FOREIGN DEPOSIT

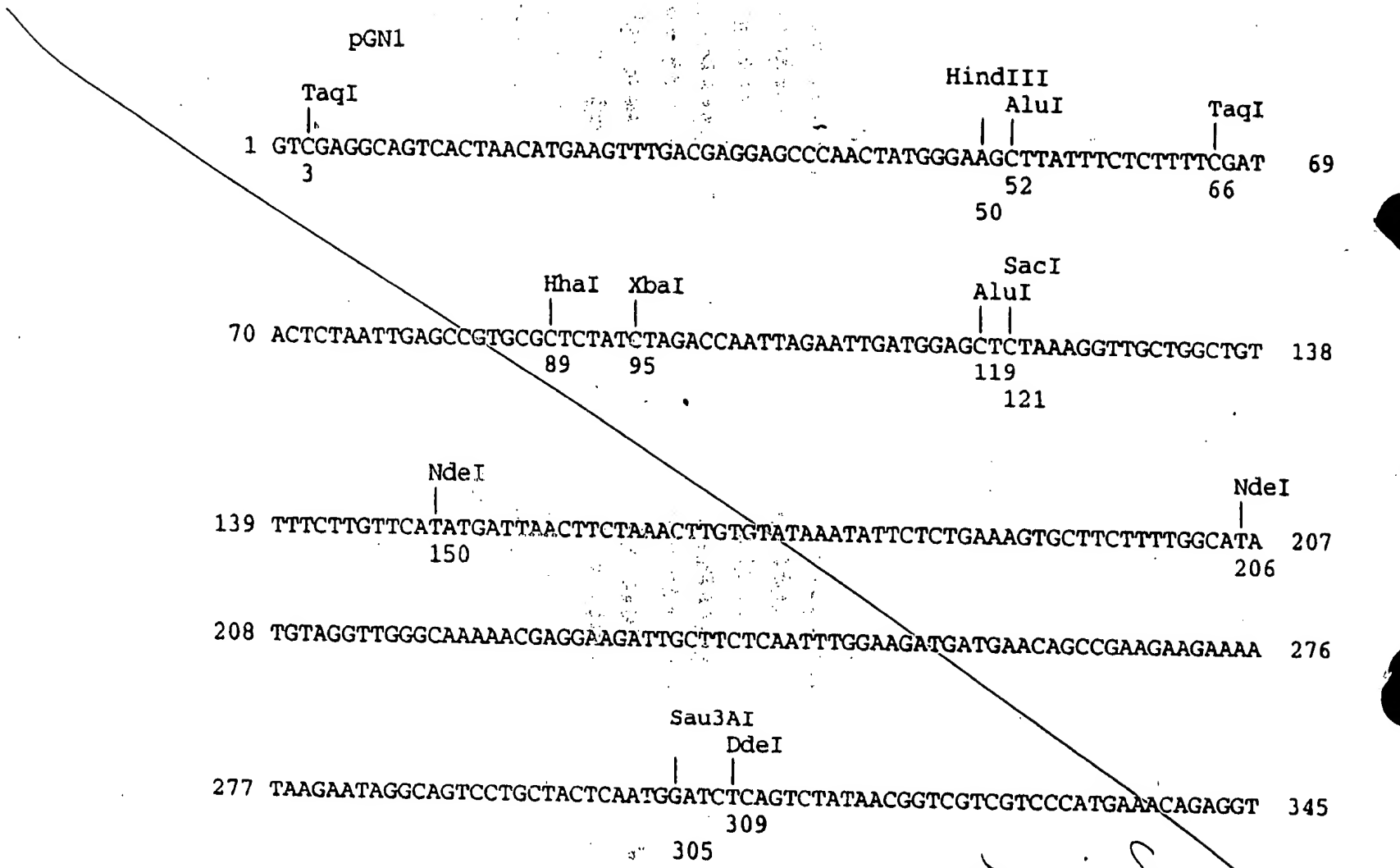


FIG. 1A

#14 paper for cancel

FOREIGN DETAILED

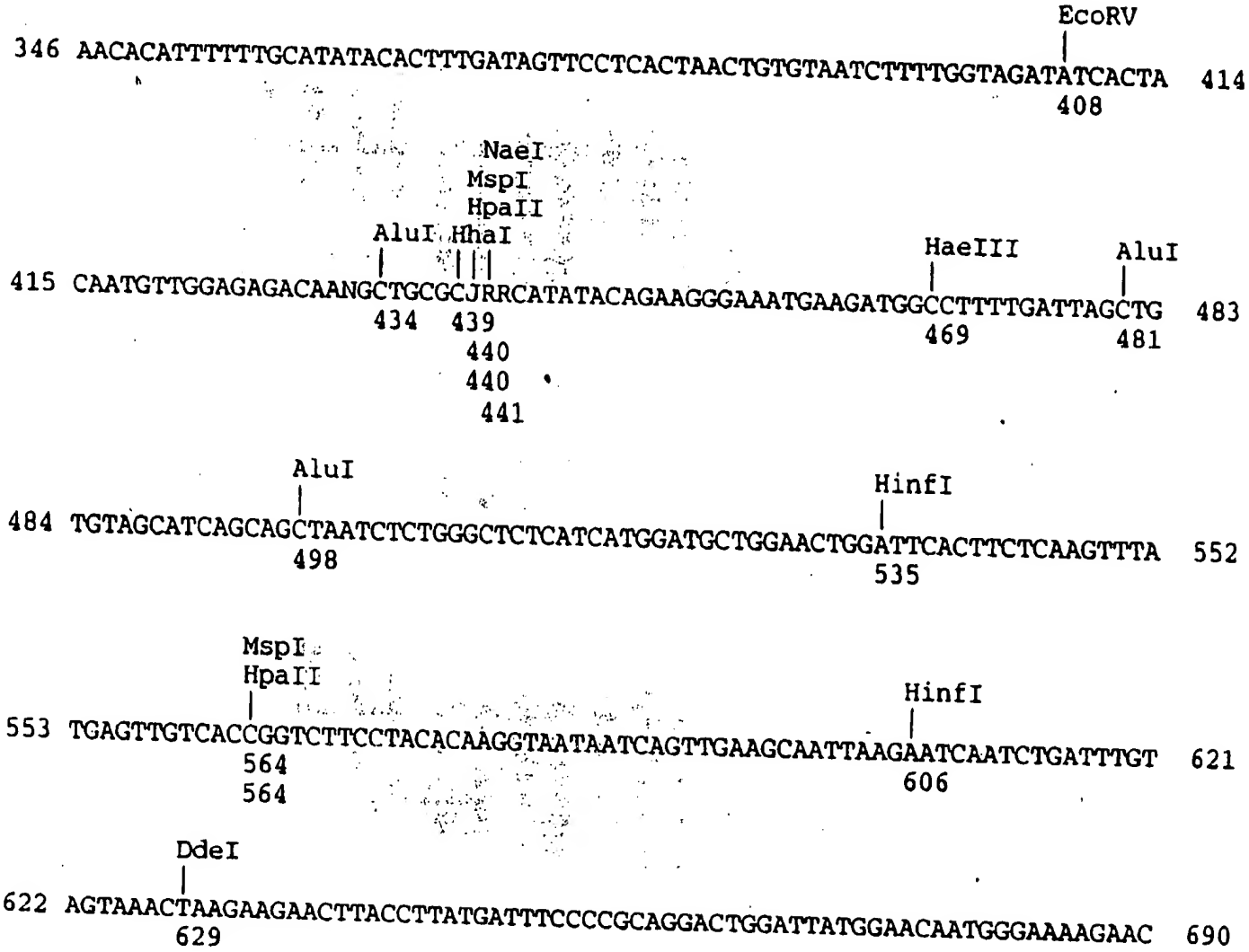


FIG. 1B

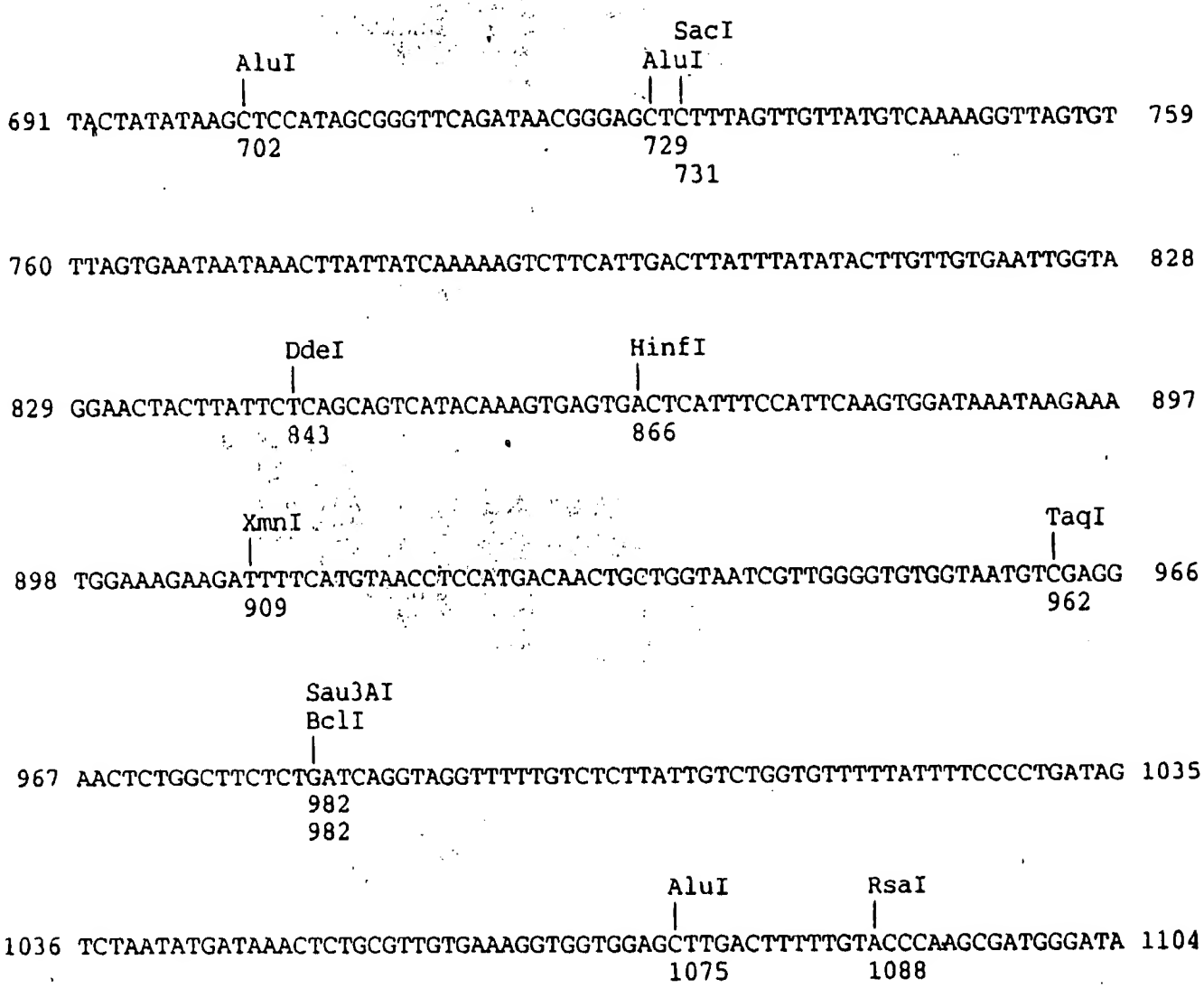

SacI
AluI
AluI
691 TACTATATAAGCTCCATAGCGGGTTCAGATAACGGGAGCTCTTTAGTTGTTATGTCAAAAGGTTAGTGT 759
702
729
731
760 TTAGTGAATAATAAACTTATTATCAAAAAGTCTTCATTGACTTATTTATATACTTGTTGTGAATTGGTA 828
DdeI
HinfI
829 GGAACTACTTATTCTCAGCAGTCATACAAAGTGAGTGACTCATTTCCATTCAAGTGGATAAATAAGAAA 897
843
866
XmnI
TaqI
898 TGGAAAGAAGATTTTCATGTAACCTCCATGACAACTGCTGGTAATCGTTGGGGTGTGGTAATGTCGAGG 966
909
962
Sau3AI
BclI
967 AACTCTGGCTTCTCTGATCAGGTAGGTTTTTGTCTCTTATTGTCTGGTGTTTTTATTTTCCCCTGATAG 1035
982
982
AluI
RsaI
1036 TCTAATATGATAAACTCTGCGTTGTGAAAGGTGGTGGAGCTTGACTTTTTGTACCCAAGCGATGGGATA 1104
1075
1088

FIG. 1C

PCT/EP97/02620

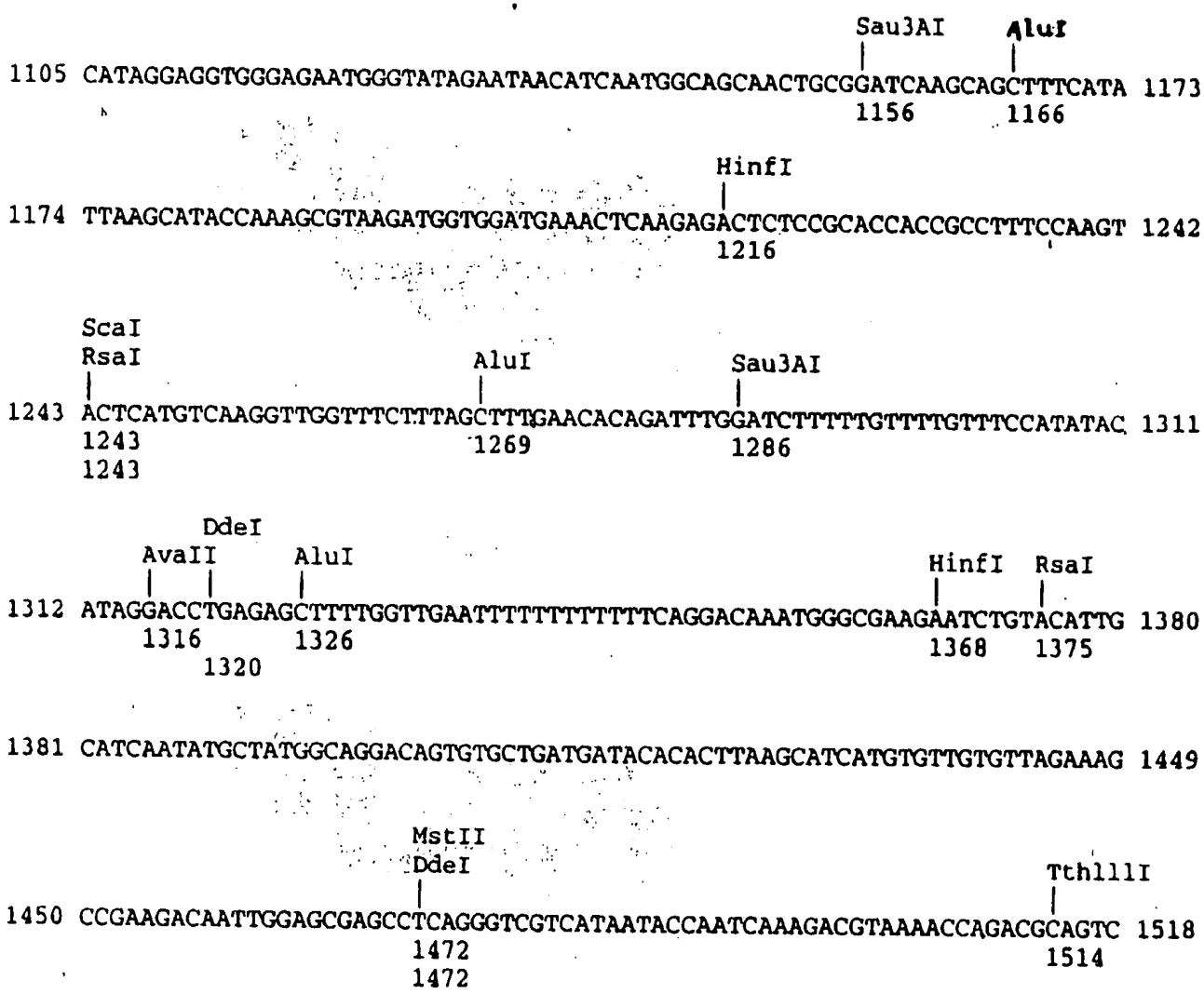


FIG. 1D

DETECTED

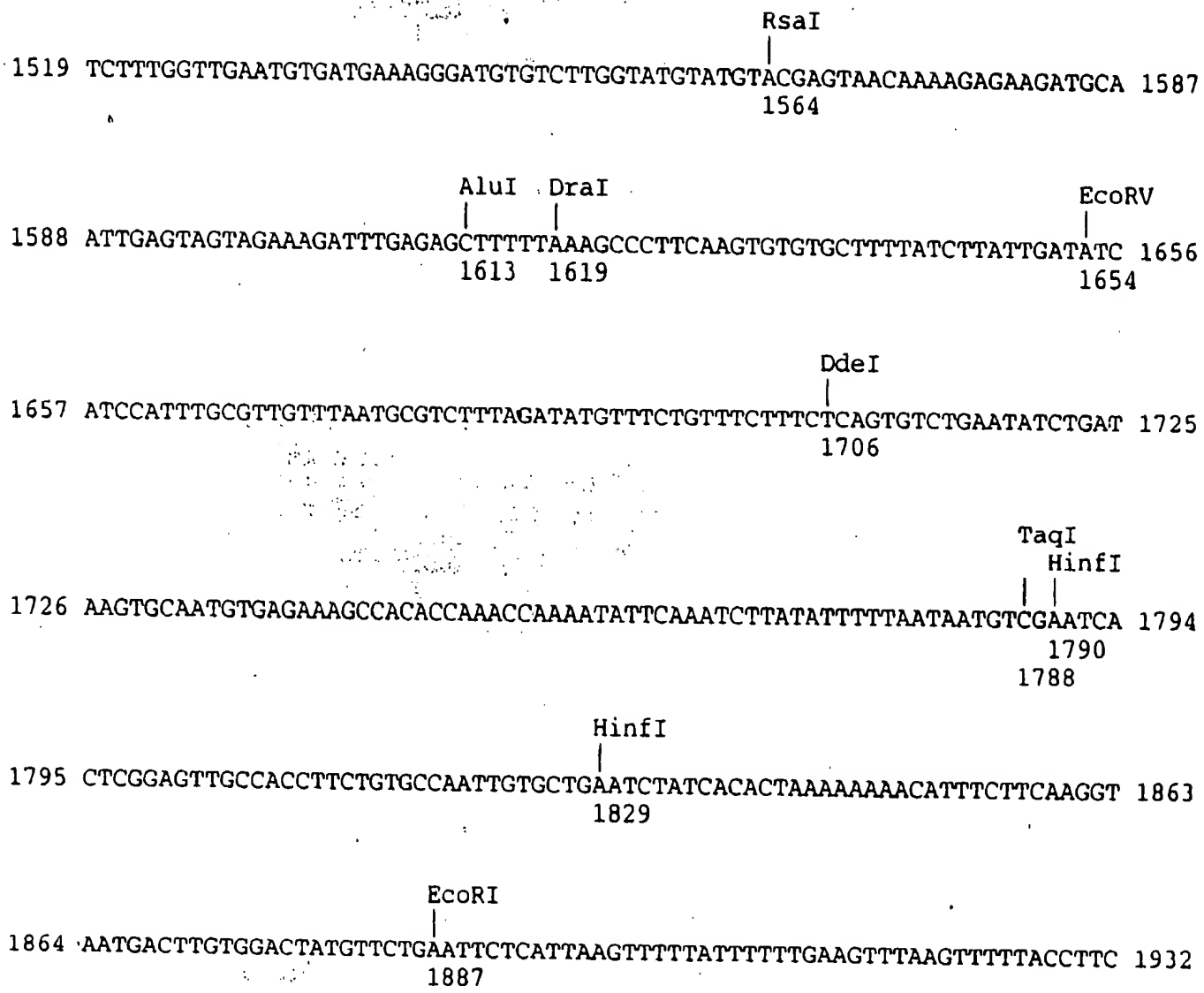


FIG. 1E

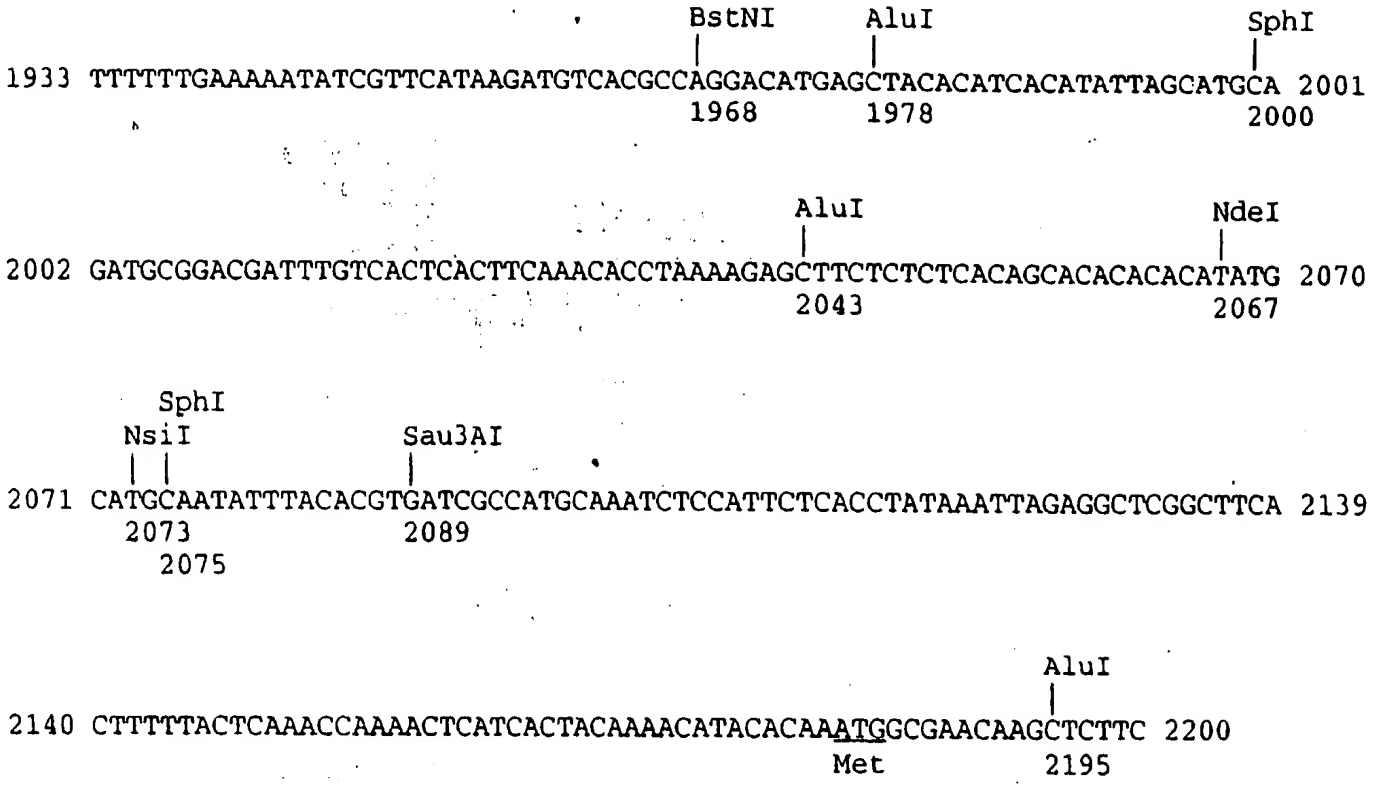

BstNI
AluI
SphI
1933 TTTTTTGAAAAATATCGTTCATAAGATGTCACGCCAGGACATGAGCTACACATCACATATTAGCATGCA 2001
1968
1978
2000
AluI
NdeI
2002 GATGCGGACGATTTGTCACTCACTTCAAACACCTAAAAGAGCTTCTCTCTCACAGCACACACACATATG 2070
2043
2067
SphI
NsiI
Sau3AI
2071 CATGCAATATTTACACGTGATCGCCATGCAAATCTCCATTCTCACCTATAAATTAGAGGCTCGGCTTCA 2139
2073
2089
2075
AluI
2140 CTTTTTACTCAAACCAAAACTCATCACTACAAAACATACACAAATGGCGAACAAGCTCTTC 2200
Met
2195

FIG. 1F

Lambda CGN1-2

NCG-186 Linear LENGTH = 4325

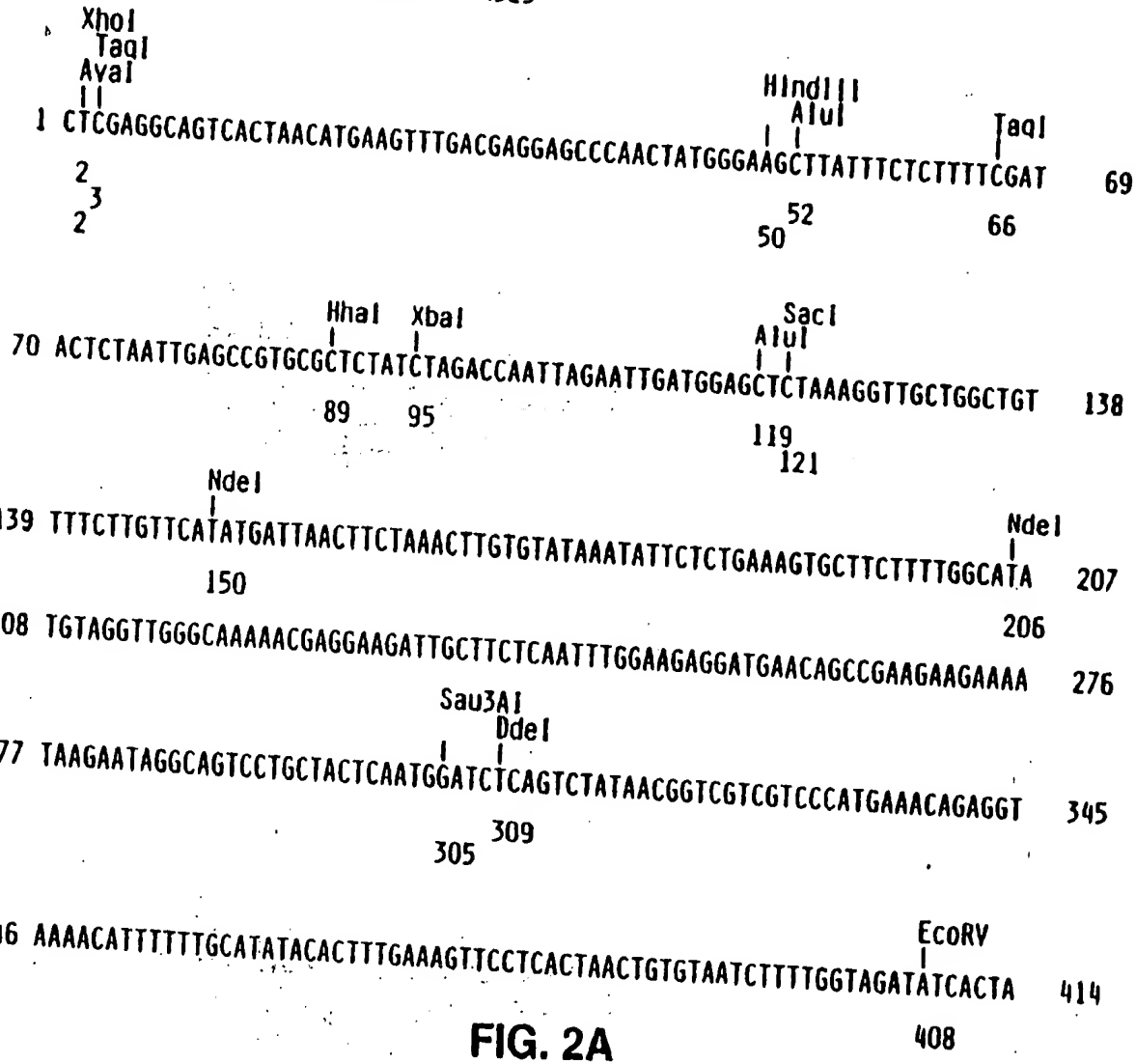
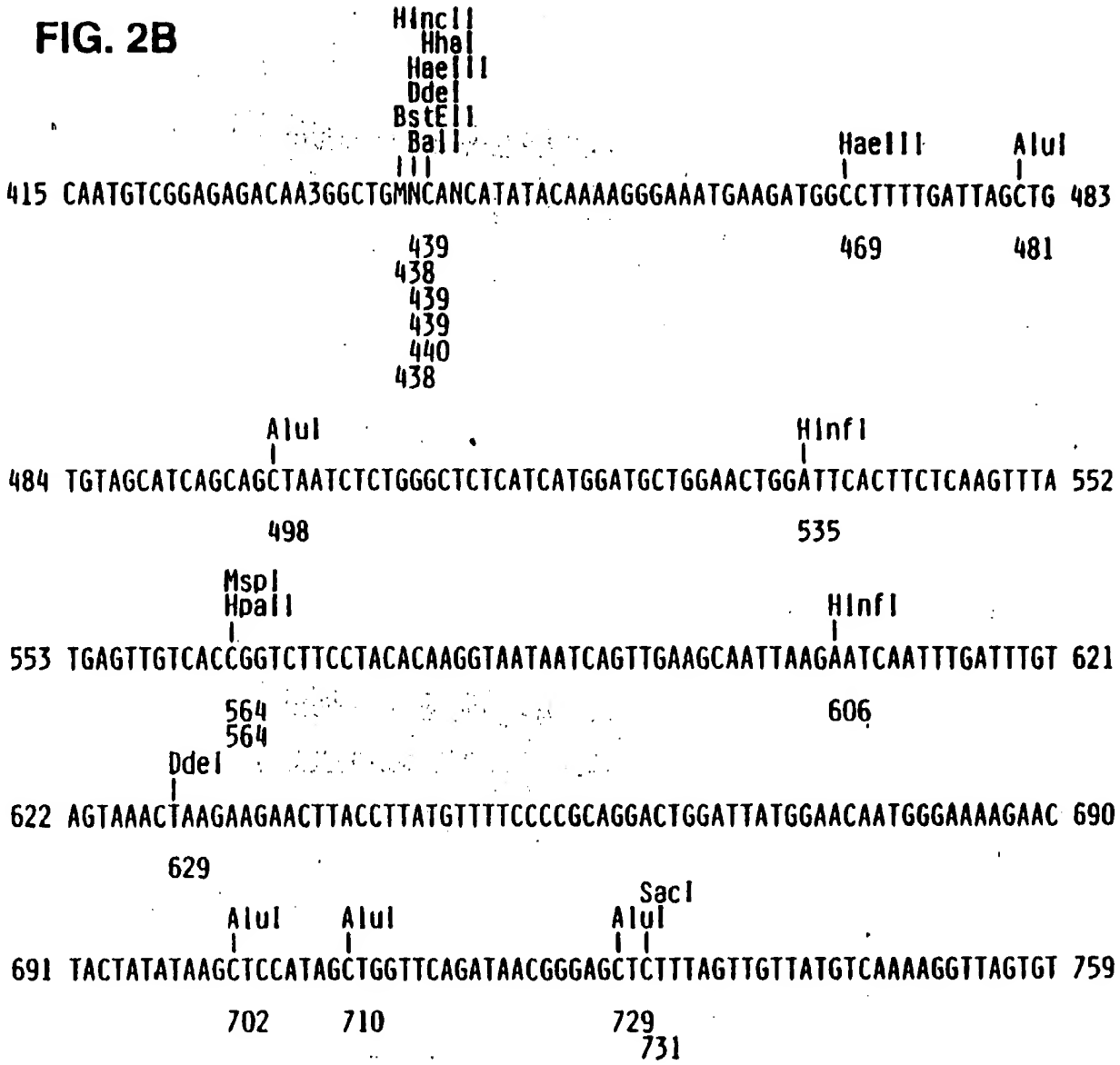


FIG. 2A

Digitized by Google

FIG. 2B



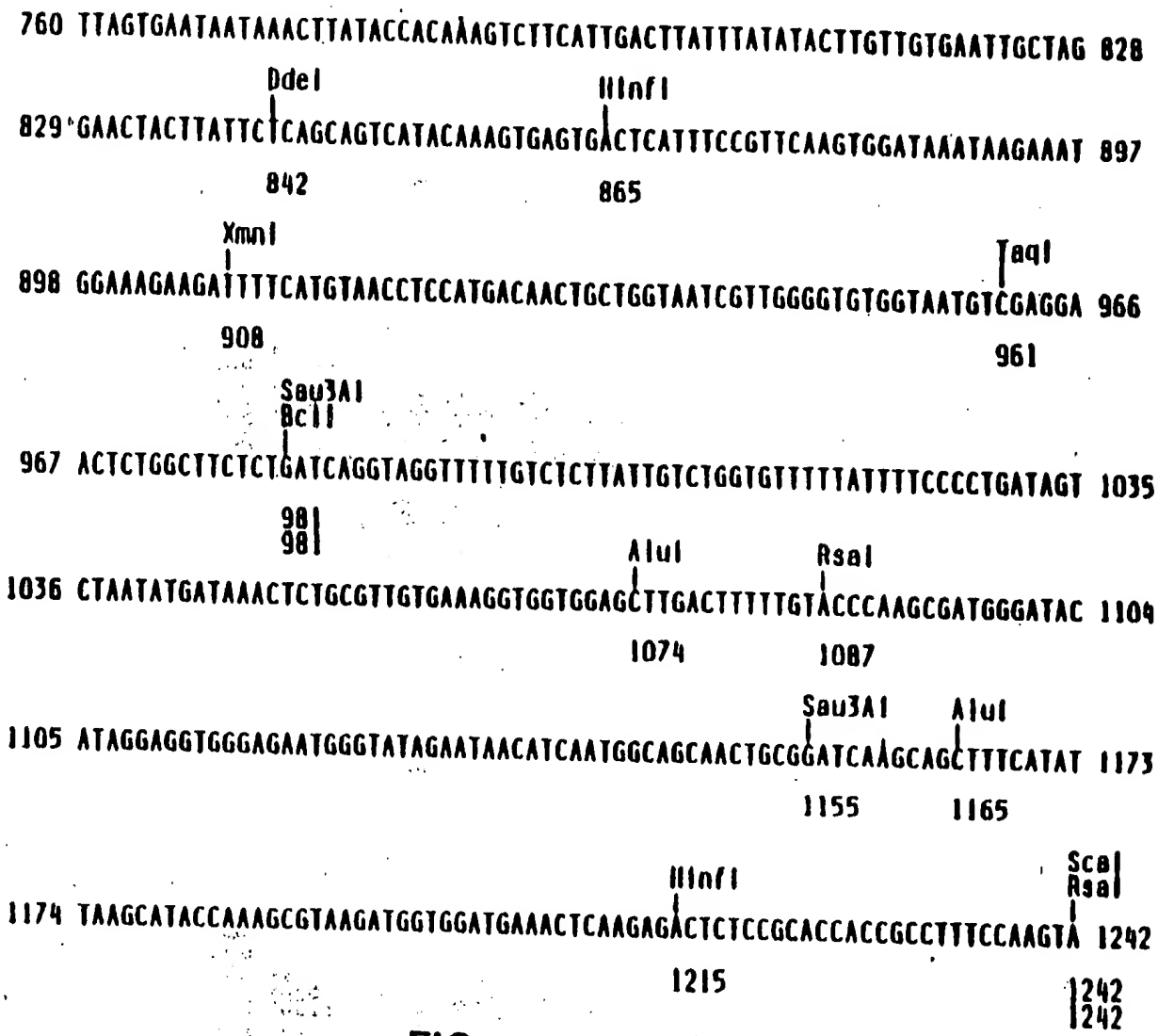

760 TTAGTGAATAATAAACTTATACCACAAAGTCTTCATTGACTTATTTATATACTTGTTGTGAATTGCTAG 828
DdeI
HinfI
829 GAACTACTTATTCTCAGCAGTCATACAAAGTGAGTGACTCATTTCCGTTCAAGTGGATAAATAAGAAAT 897
842
865
XmnI
TaqI
898 GGAAAGAAGATTTTCATGTAACCTCCATGACAACTGCTGGTAATCGTTGGGGTGTGGTAATGTCGAGGA 966
908
961
Sau3AI
BclI
967 ACTCTGGCTTCTCTGATCAGGTAGGTTTTTGTCTCTTATTGTCTGGTGTTTTTATTTTCCCCTGATAGT 1035
98
98
AluI
RsaI
1036 CTAATATGATAAACTCTGCGTTGTGAAAGGTGGTGGAGCTTGACTTTTTGTACCCAAGCGATGGGATAC 1104
1074
1087
Sau3AI
AluI
1105 ATAGGAGGTGGGAGAATGGGTATAGAATAACATCAATGGCAGCAACTGCGGATCAAGCAGCTTTCATAT 1173
1155
1165
HinfI
ScaI
RsaI
1174 TAAGCATACCAAAGCGTAAGATGGTGGATGAAACTCAAGAGACTCTCCGCACCACCGCCTTTCCAAGTA 1242
1215
1242
1242

FIG. 2C

FOREIGN DETAILED

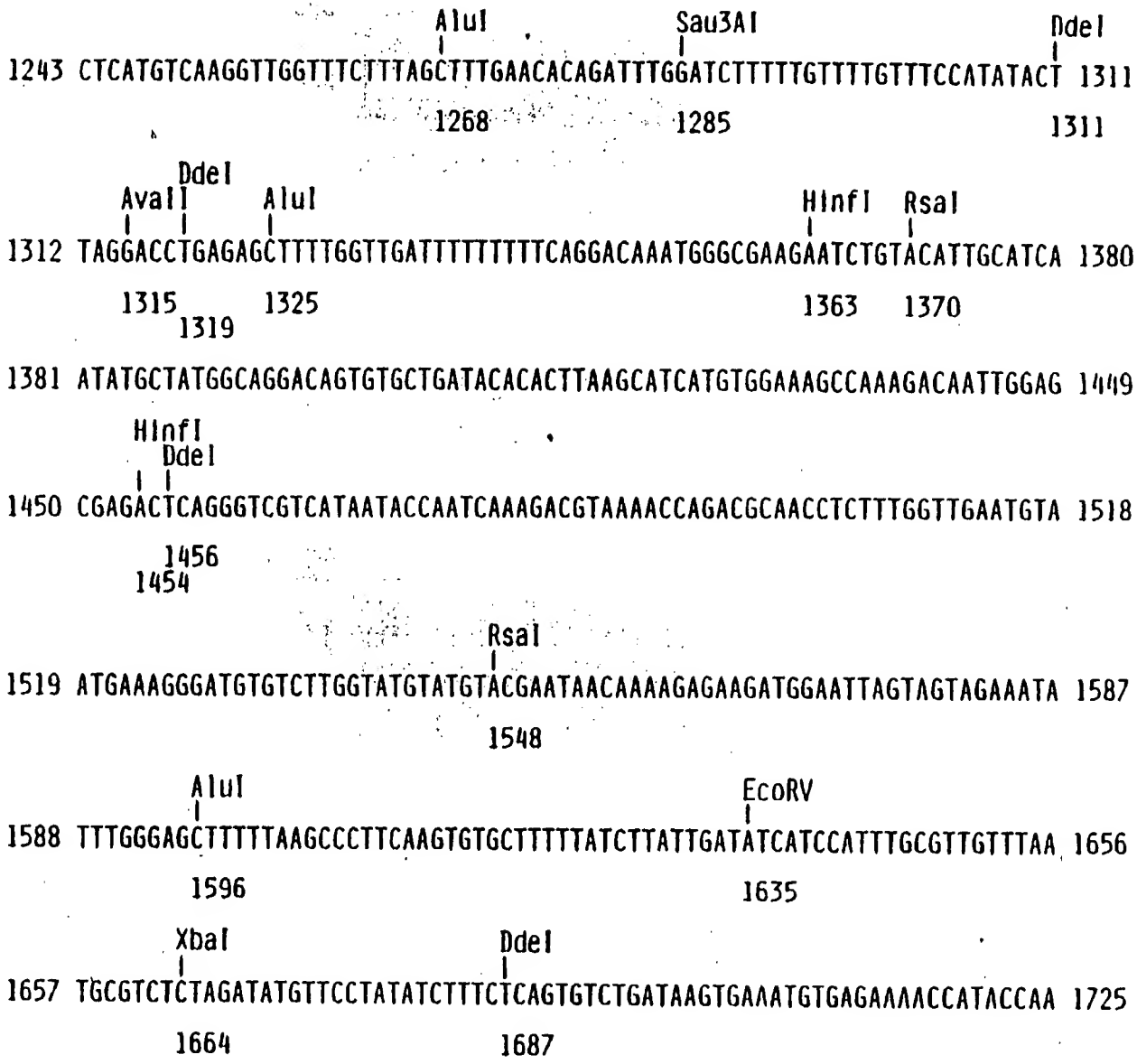


FIG. 2D

TOSTED DEFEATED

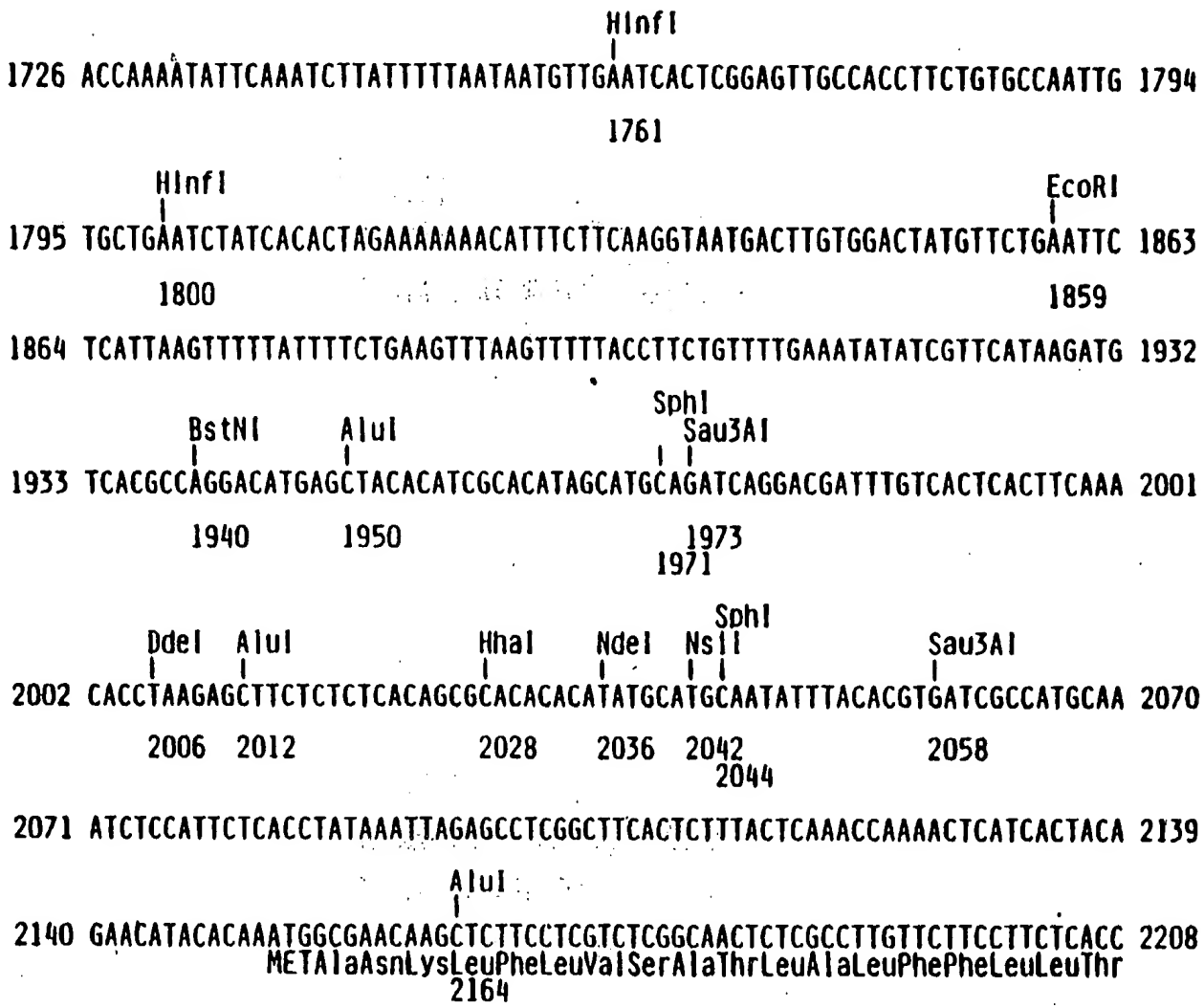


FIG. 2E

PCT/EP2001/013260

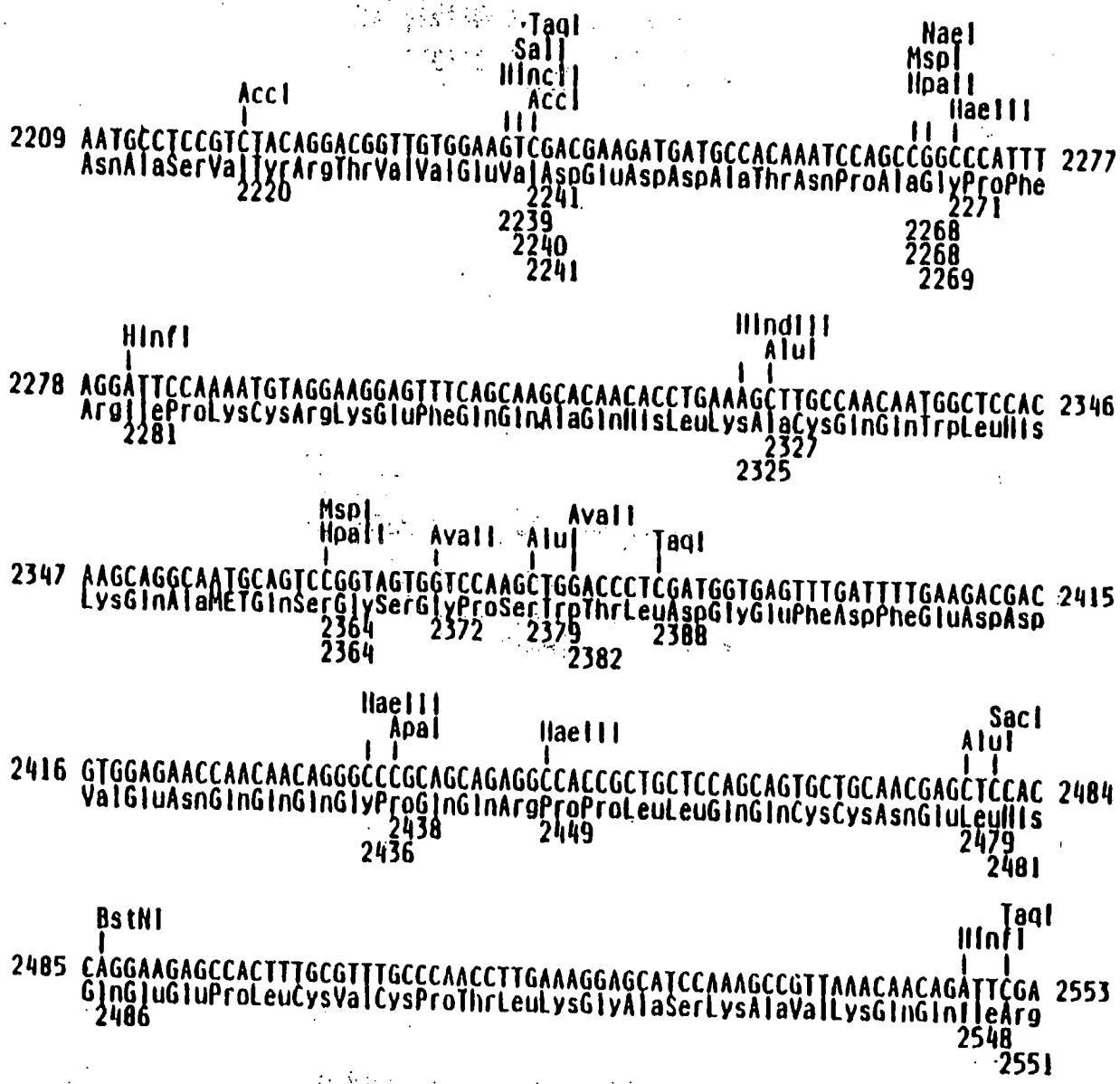


FIG. 2F

DOCKET NO. 26260

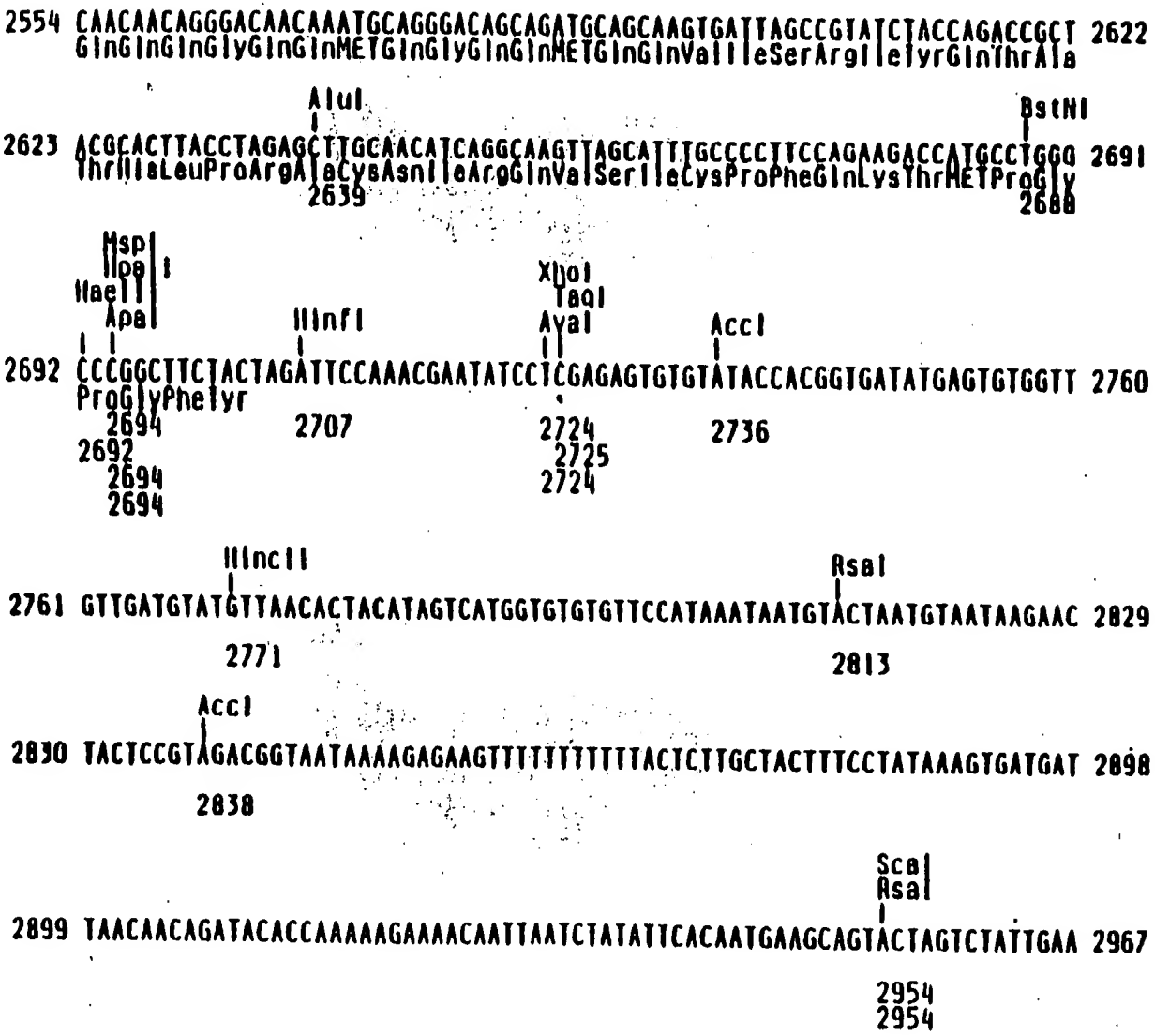


FIG. 2G

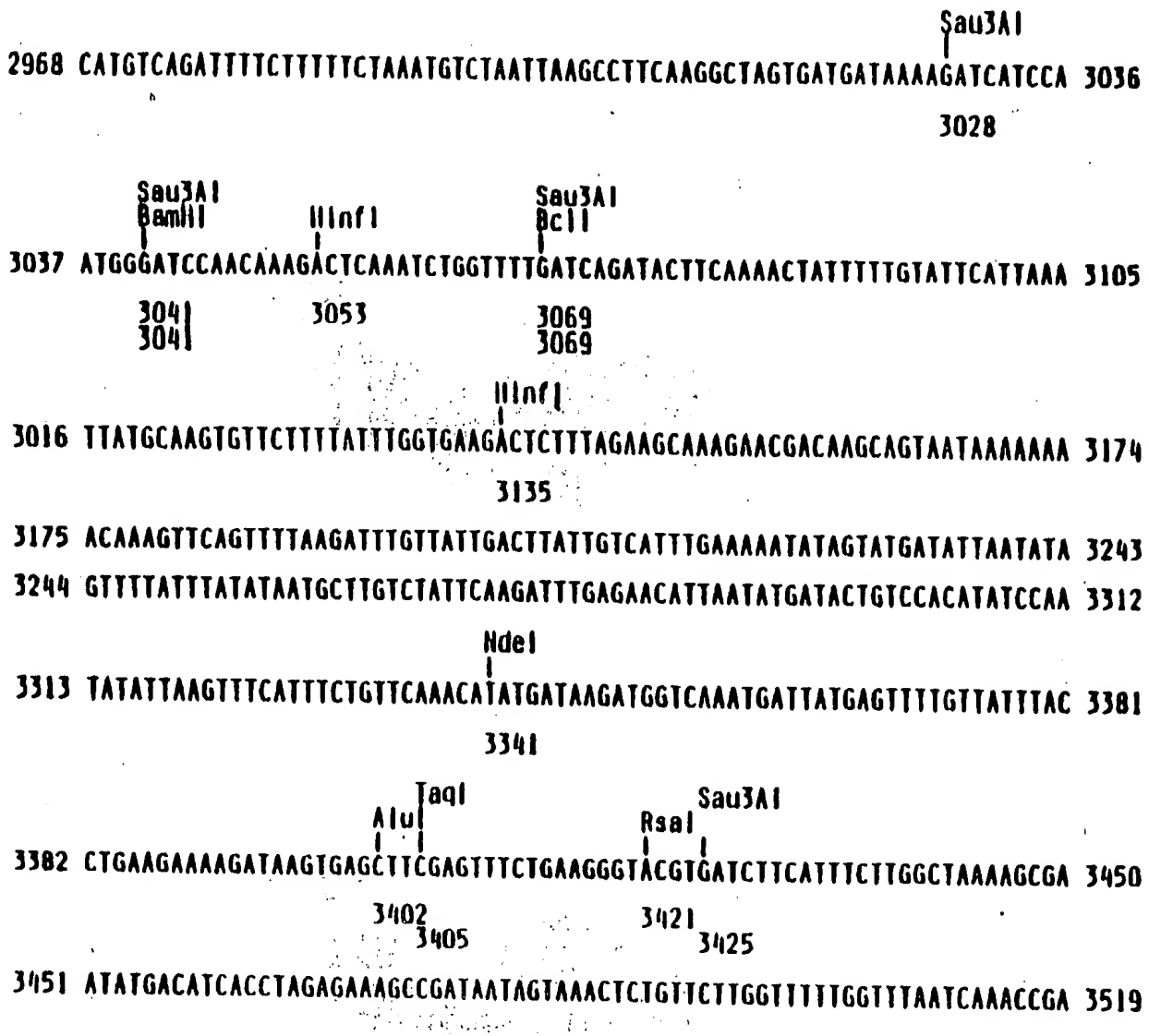

Sau3AI
2968 CATGTCAGATTTTCTTTTTCTAAATGTCTAATTAAGCCTTCAAGGCTAGTGATGATAAAAGATCATCCA 3036
3028
Sau3AI
BamHI
HInf I
Sau3AI
BclI
3037 ATGGGATCCAACAAAGACTCAAATCTGGTTTTGATCAGATACTTCAAAACTATTTTTGTATTCATTAAA 3105
3041
3041
3053
3069
3069
HInf I
3016 TTATGCAAGTGTTCTTTTATTTGGTGAAGACTCTTTAGAAGCAAAGAACGACAAGCAGTAATAAAAAAA 3174
3135
3175 ACAAAGTTCAGTTTTAAGATTTGTTATTGACTTATTGTCATTTGAAAAATATAGTATGATATTAATATA 3243
3244 GTTTTATTTATATAATGCTTGTCTATTCAAGATTTGAGAACATTAATATGATACTGTCCACATATCCAA 3312
NdeI
3313 TATATTAAGTTTCATTTCTGTTCAAACATATGATAAGATGGTCAAATGATTATGAGTTTTGTTATTTAC 3381
3341
TaqI
Sau3AI
AluI
RsaI
3382 CTGAAGAAAAGATAAGTGAGCTTCGAGTTTCTGAAGGGTACGTGATCTTCATTTCTTGGCTAAAAGCGA 3450
3402
3405
3421
3425
3451 ATATGACATCACCTAGAGAAAGCCGATAATAGTAAACTCTGTTCTTGGTTTTTGGTTTAATCAAACCGA 3519

FIG. 2H

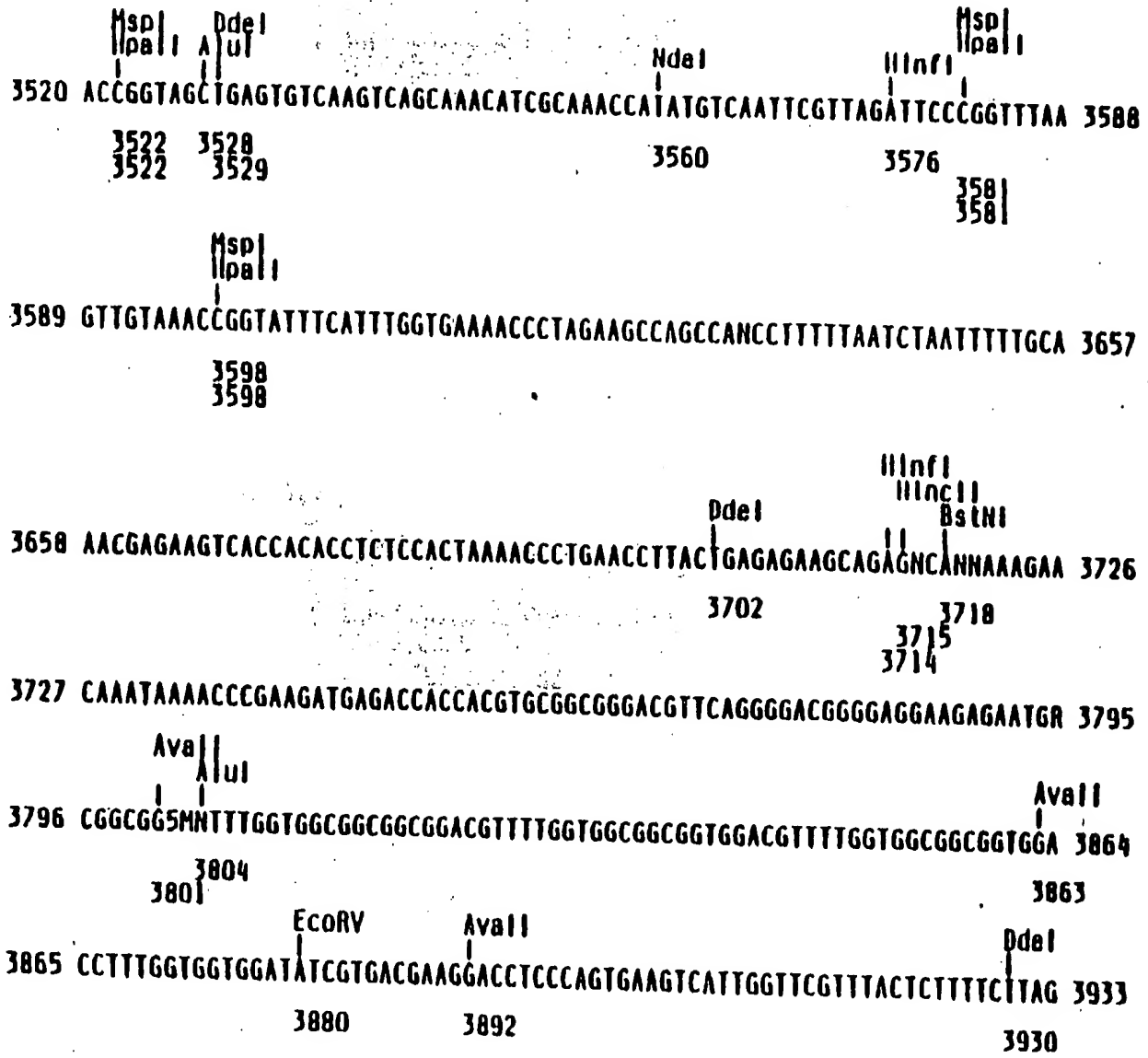

MspI
pa
AluI
DdeI
NdeI
IIInfI
MspI
pa
3520 ACCGGTAGCTGAGTGTCAAGTCAGCAAACATCGCAAACCATATGTCAATTCGTTAGATTCCCGGTTTAA 3588
3522
3522
3528
3529
3560
3576
358
358
MspI
pa
3589 GTTGTAAACCGGTATTTCATTTGGTGAAAACCCTAGAAGCCAGCCANCCTTTTTAATCTAATTTTTGCA 3657
3598
3598
IIInfI
IIIncII
DdeI
BstNI
3658 AACGAGAAGTCACCACACCTCTCCACTAAAACCCTGAACCTTACTGAGAGAAGCAGAGNCANNAAAGAA 3726
3702
3718
3715
3714
3727 CAAATAAAACCCGAAGATGAGACCACCACGTGCGGCGGGACGTTCAGGGGACGGGGAGGAAGAGAATGR 3795
AvaI
AluI
AvaII
3796 CGGCGGSMNTTTGGTGGCGGCGGCGGACGTTTTGGTGGCGGCGGTGGACGTTTTGGTGGCGGCGGTGGA 3864
3801
3804
3863
EcoRV
AvaII
DdeI
3865 CCTTTGGTGGTGGATATCGTGACGAAGGACCTCCCAGTGAAGTCATTGGTTCGTTTACTCTTTTCTTAG 3933
3880
3892
3930

FIG. 2I

PCT/EP92/02200

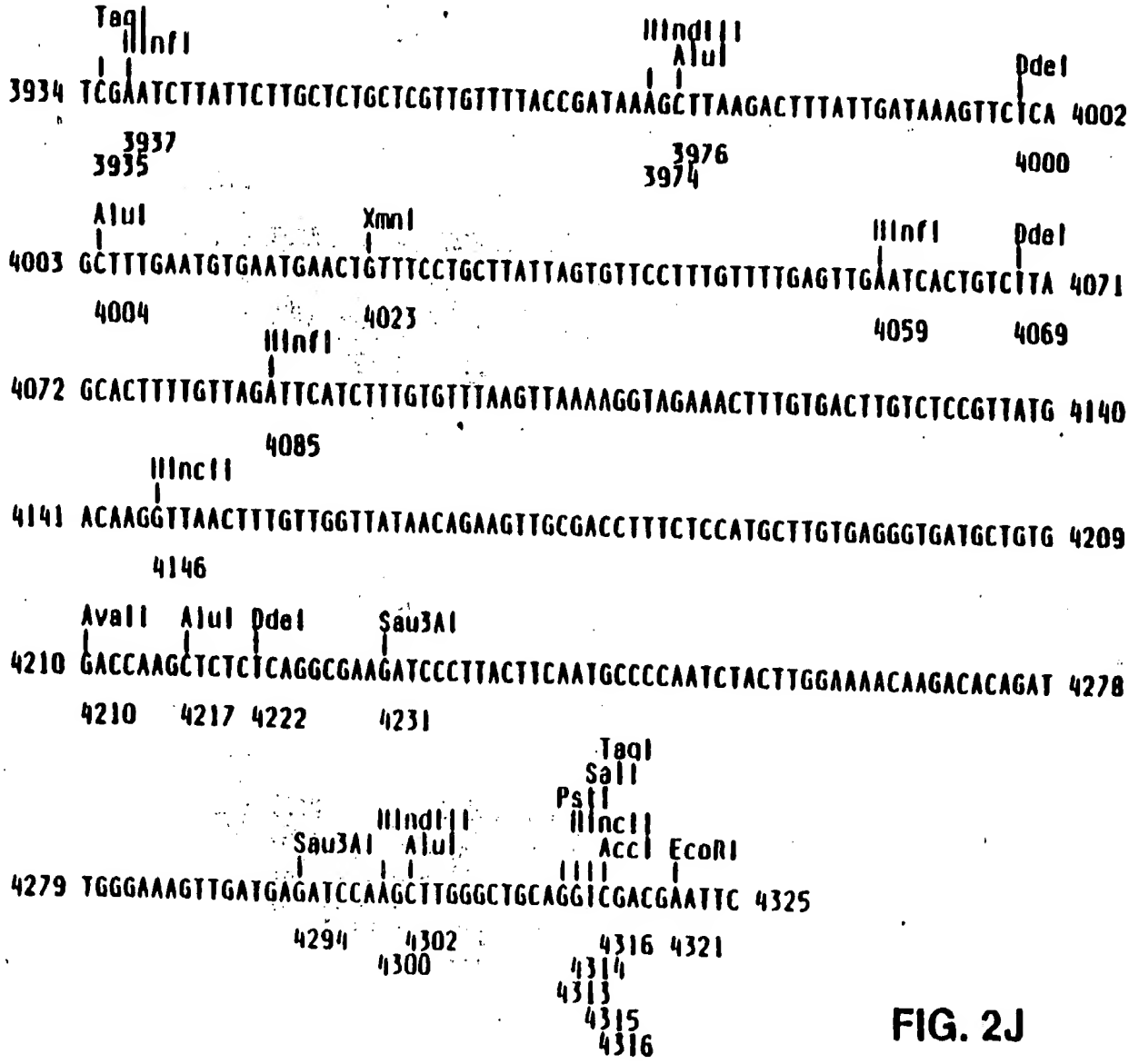


FIG. 2J

PCT/EP97/06760

Brassica campestris ACP Genomic Sequence

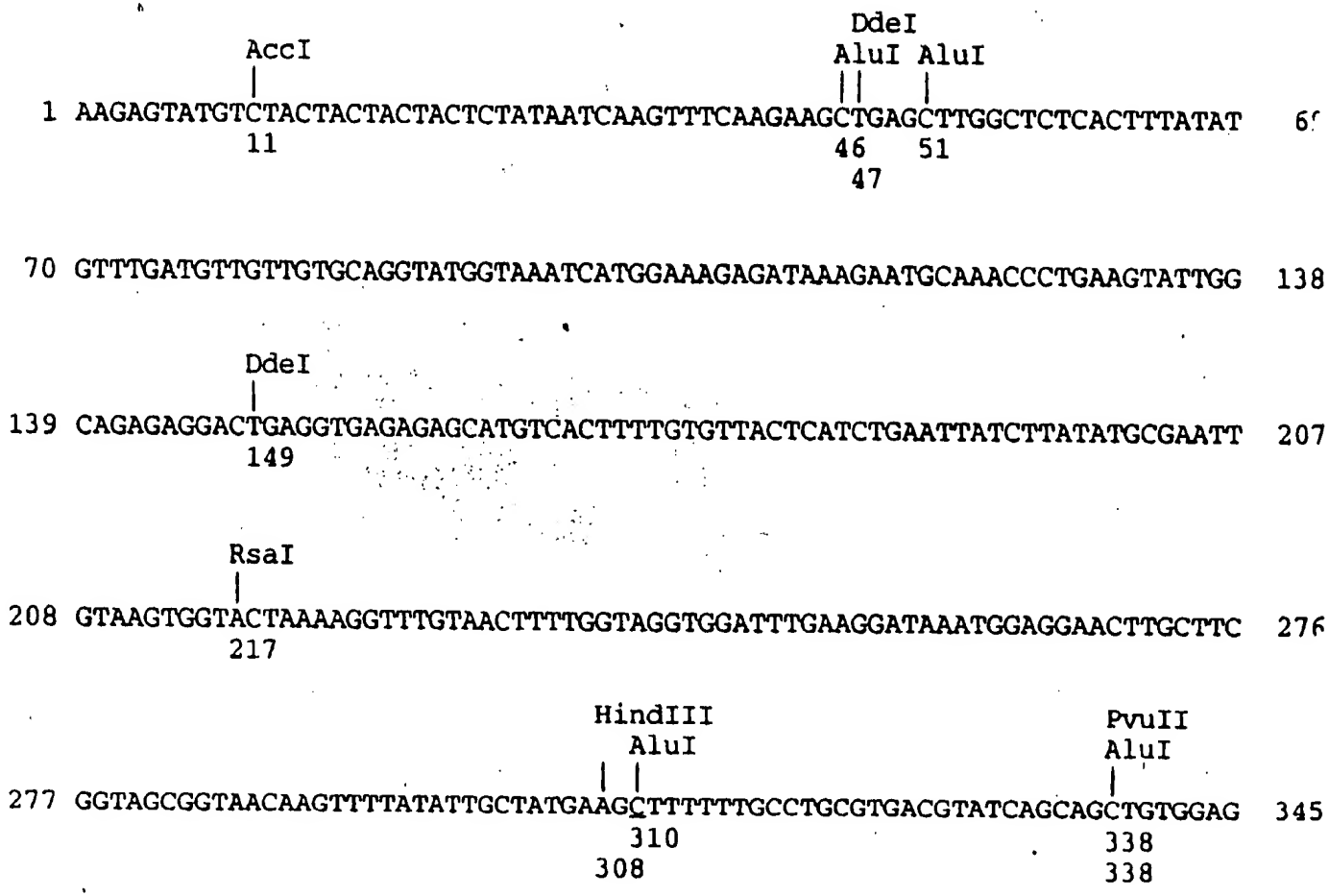


FIG. 3A

TOP SECRET DEFENSE

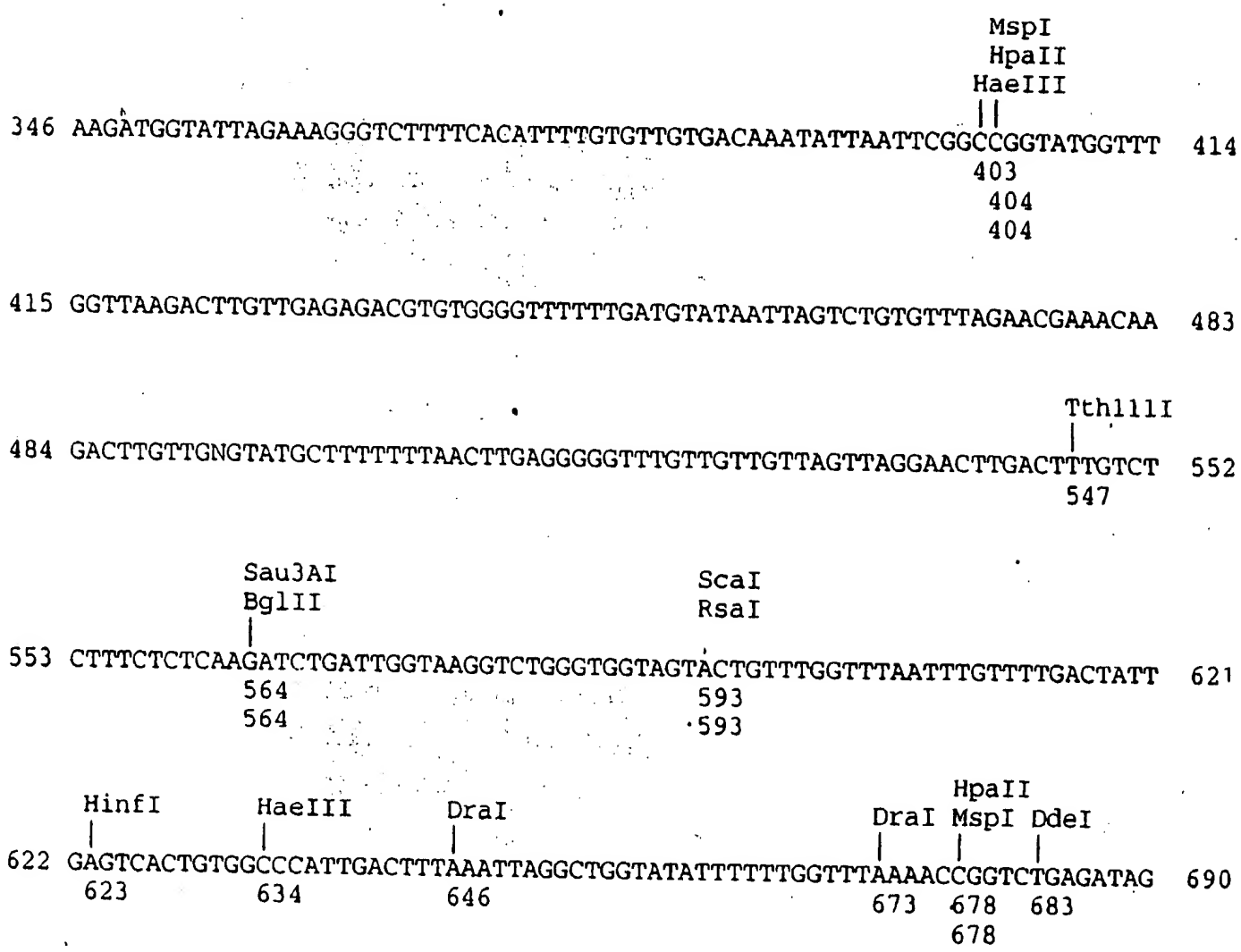


FIG. 3B

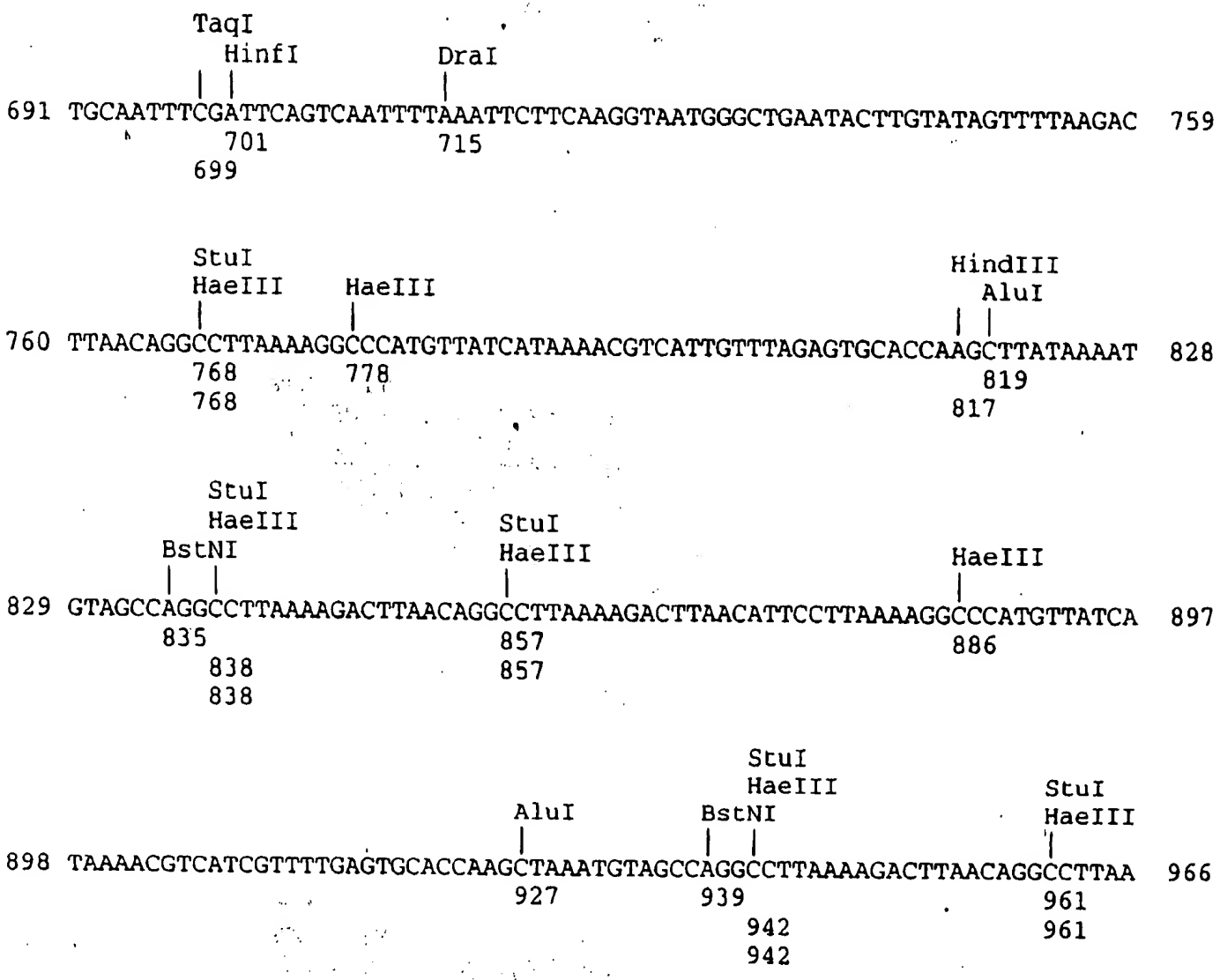

TaqI
HinfI
DraI
691 TGCAATTTCGATTCAGTCAATTTTAAATTCTTCAAGGTAATGGGCTGAATACTTGTATAGTTTTAAGAC 759
701
715
699
StuI
HaeIII
HaeIII
HindIII
AluI
760 TTAACAGGCCTTAAAAGGCCCATGTTATCATAAAACGTCATTGTTTAGAGTGCACCAAGCTTATAAAAT 828
768
778
819
768
817
StuI
HaeIII
BstNI
StuI
HaeIII
HaeIII
829 GTAGCCAGGCCTTAAAAGACTTAACAGGCCTTAAAAGACTTAACATTCCTTAAAAGGCCCATGTTATCA 897
835
857
886
838
857
838
StuI
HaeIII
AluI
BstNI
StuI
HaeIII
898 TAAAACGTCATCGTTTTGAGTGCACCAAGCTAAATGTAGCCAGGCCTTAAAAGACTTAACAGGCCTTAA 966
927
939
961
942
961
942

FIG. 3C

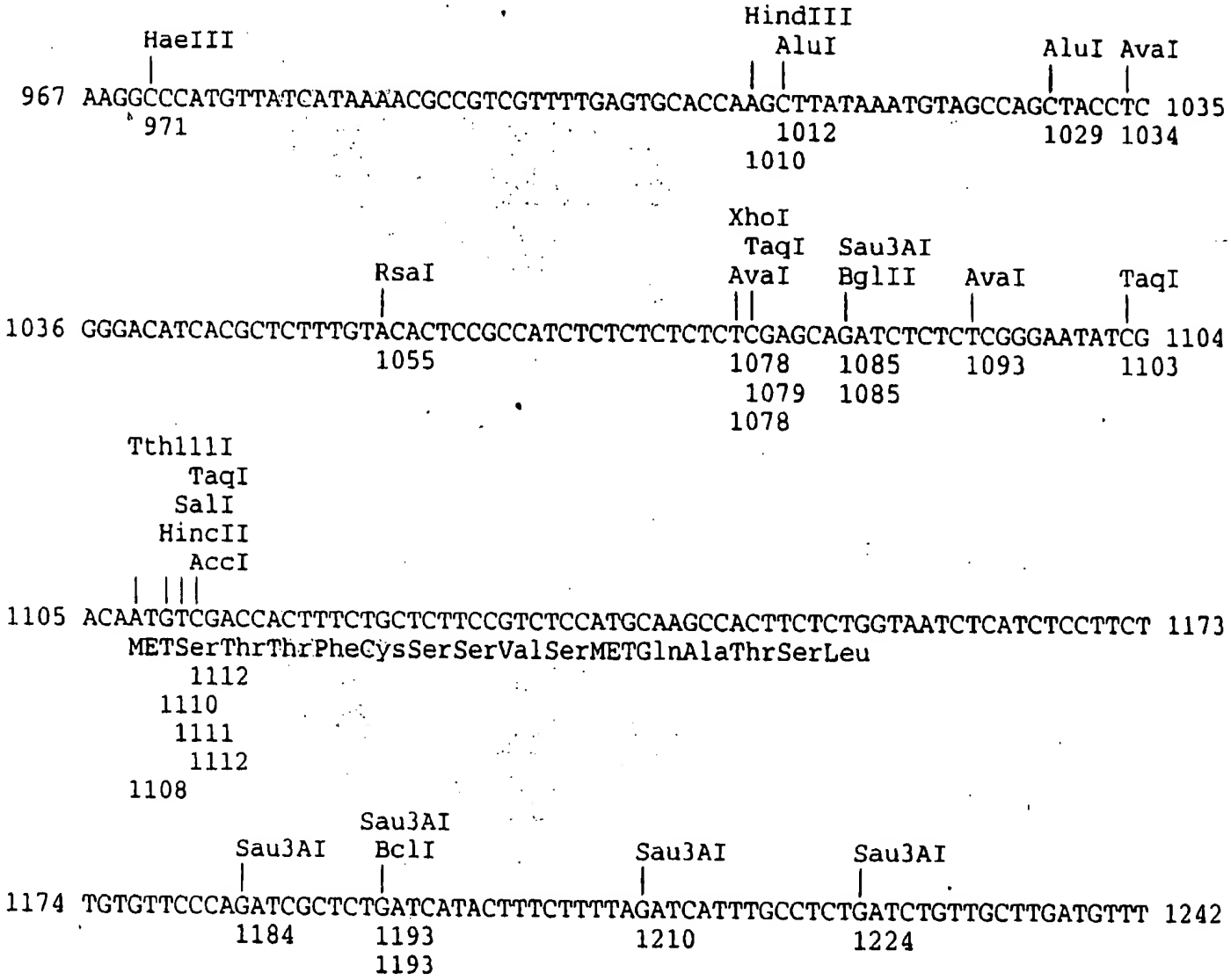

HaeIII
HindIII
AluI
AluI AvaI
967 AAGGCCCATGTTATCATAAAACGCCGTCGTTTTGAGTGCACCAAGCTTATAAATGTAGCCAGCTACCTC 1035
971
1012
1029 1034
1010
XhoI
TaqI Sau3AI
RsaI
AvaI BglII AvaI TaqI
1036 GGGACATCACGCTCTTTGTACACTCCGCCATCTCTCTCTCTCTCGAGCAGATCTCTCTCGGGAATATCG 1104
1055
1078 1085 1093 1103
1079 1085
1078
Tth111I
TaqI
SalI
HincII
AccI
1105 ACAATGTCGACCACTTTCTGCTCTTCCGTCTCCATGCAAGCCACTTCTCTGGTAATCTCATCTCCTTCT 1173
METSerThrThrPheCysSerSerValSerMETGlnAlaThrSerLeu
1112
1110
1111
1112
1108
Sau3AI
Sau3AI BclI
Sau3AI
Sau3AI
1174 TGTGTTCCCAGATCGCTCTGATCATACTTTCTTTTAGATCATTTGCCTCTGATCTGTTGCTTGATGTTT 1242
1184
1193
1210
1224
1193

FIG. 3D

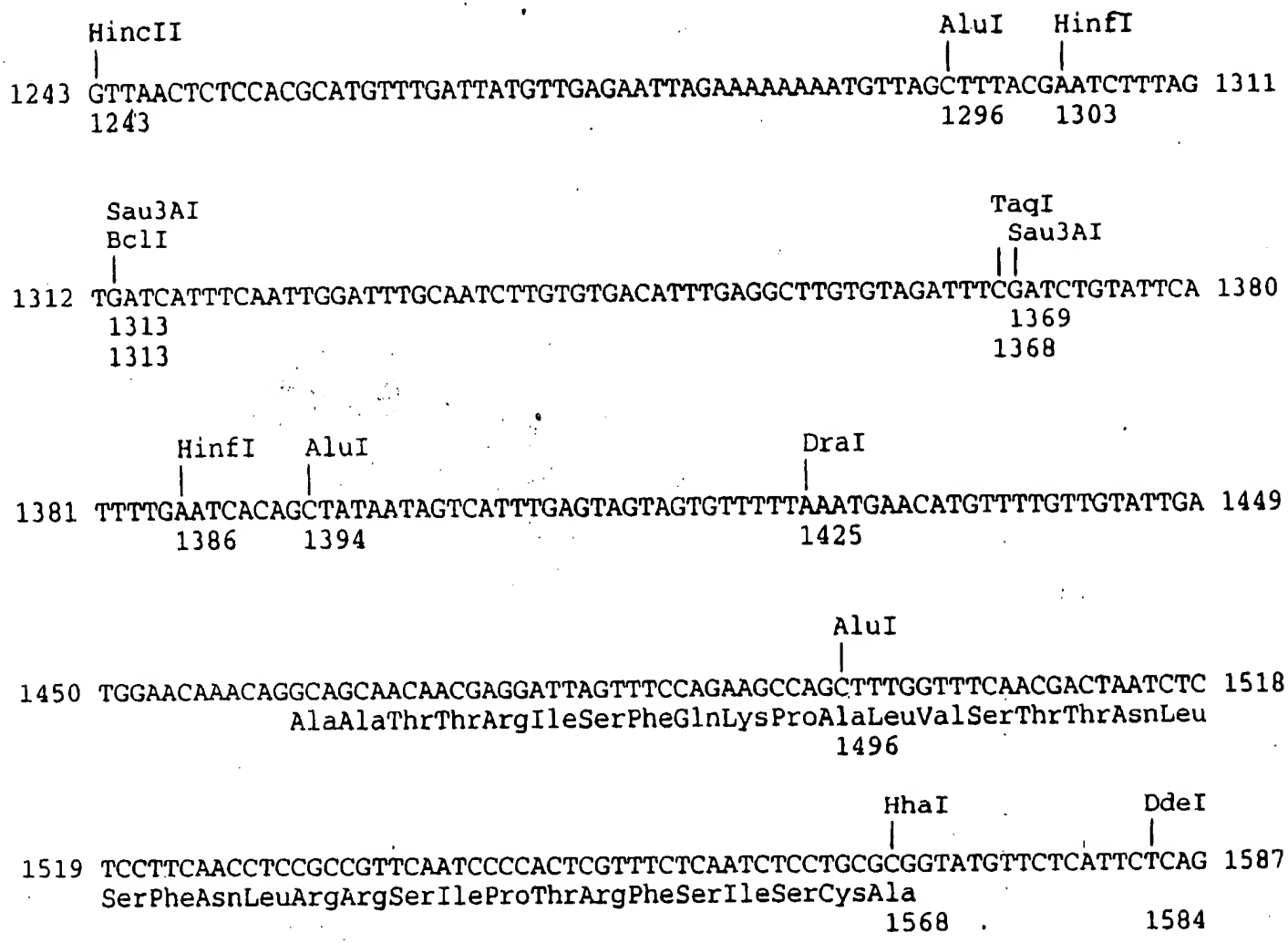

HincII
AluI
HinfI
1243 GTTAACTCTCCACGCATGTTTGATTATGTTGAGAATTAGAAAAAAAATGTTAGCTTTACGAATCTTTAG 1311
1243
1296
1303
Sau3AI
BclI
TaqI
Sau3AI
1312 TGATCATTTCAATTGGATTTGCAATCTTGTGTGACATTTGAGGCTTGTGTAGATTTCGATCTGTATTCA 1380
1313
1313
1369
1368
HinfI
AluI
DraI
1381 TTTTGAATCACAGCTATAATAGTCATTTGAGTAGTAGTGTTTTTAAATGAACATGTTTTGTTGTATTGA 1449
1386
1394
1425
AluI
1450 TGGAACAAACAGGCAGCAACAACGAGGATTAGTTTCCAGAAGCCAGCTTTGGTTTCAACGACTAATCTC 1518
AlaAlaThrThrArgIleSerPheGlnLysProAlaLeuValSerThrThrAsnLeu
1496
HhaI
DdeI
1519 TCCTTCAACCTCCGCCGTTCAATCCCCACTCGTTTCTCAATCTCCTGCGCGGTATGTTCTCATTCTCAG 1587
SerPheAsnLeuArgArgSerIleProThrArgPheSerIleSerCysAla
1568
1584

FIG. 3E

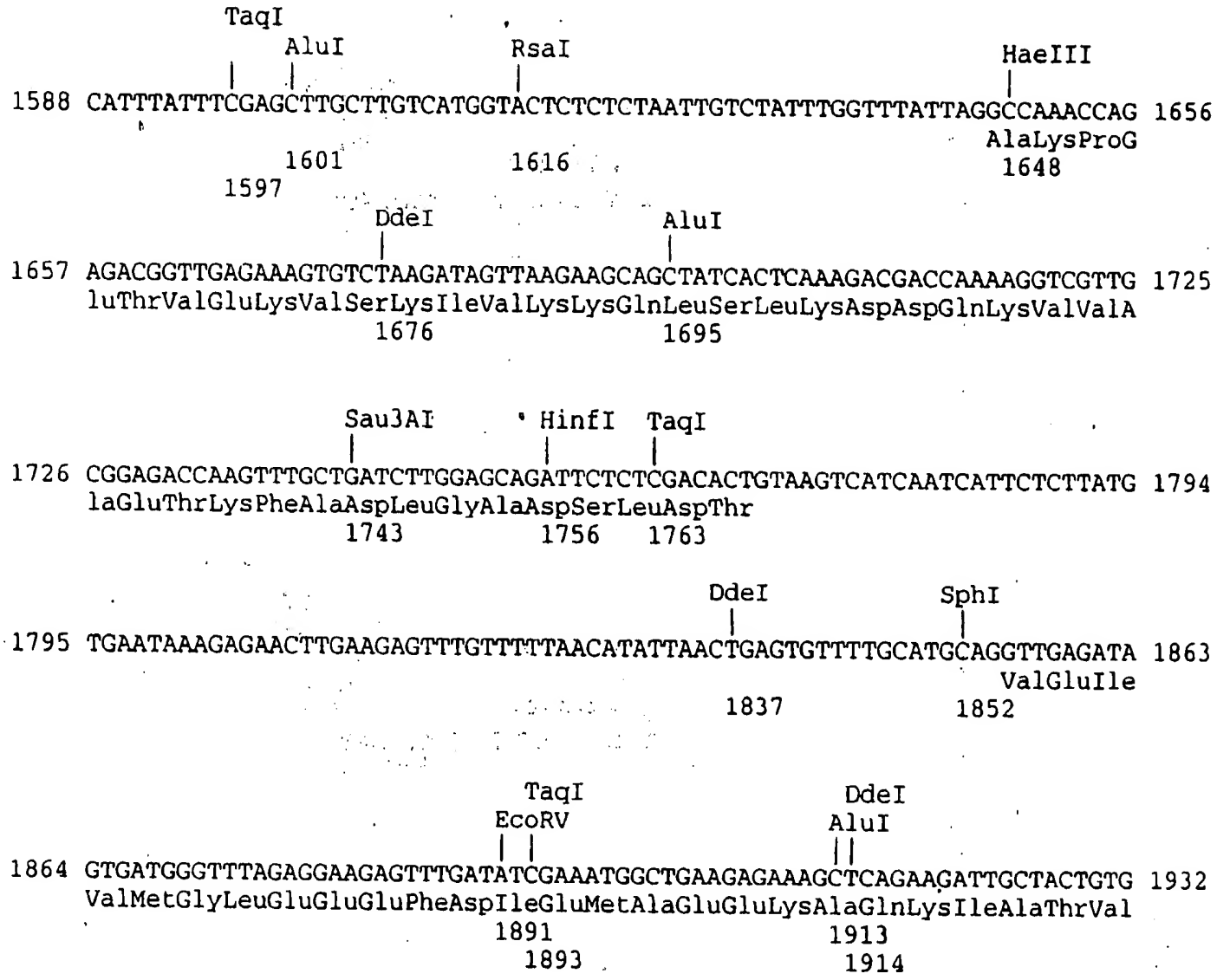

TaqI
AluI
RsaI
HaeIII
1588 CATTTATTTCGAGCTTGCTTGTCATGGTACTCTCTCTAATTGTCTATTTGGTTTATTAGGCCAAACCAG 1656
AlaLysProG
1601
1616
1648
1597
DdeI
AluI
1657 AGACGGTTGAGAAAGTGTCTAAGATAGTTAAGAAGCAGCTATCACTCAAAGACGACCAAAAGGTCGTTG 1725
luThrValGluLysValSerLysIleValLysLysGlnLeuSerLeuLysAspAspGlnLysValValA
1676
1695
Sau3AI
HinfI
TaqI
1726 CGGAGACCAAGTTTGCTGATCTTGGAGCAGATTCTCTCGACACTGTAAGTCATCAATCATTCTCTTATG 1794
laGluThrLysPheAlaAspLeuGlyAlaAspSerLeuAspThr
1743
1756
1763
DdeI
SphI
1795 TGAATAAAGAGAACTTGAAGAGTTTGTTTTTAACATATTAACTGAGTGTTTTGCATGCAGGTTGAGATA 1863
ValGluIle
1837
1852
TaqI
EcoRV
DdeI
AluI
1864 GTGATGGGTTTAGAGGAAGAGTTTGATATCGAAATGGCTGAAGAGAAAGCTCAGAAGATTGCTACTGTG 1932
ValMetGlyLeuGluGluGluPheAspIleGluMetAlaGluGluLysAlaGlnLysIleAlaThrVal
1891
1893
1913
1914

FIG. 3F

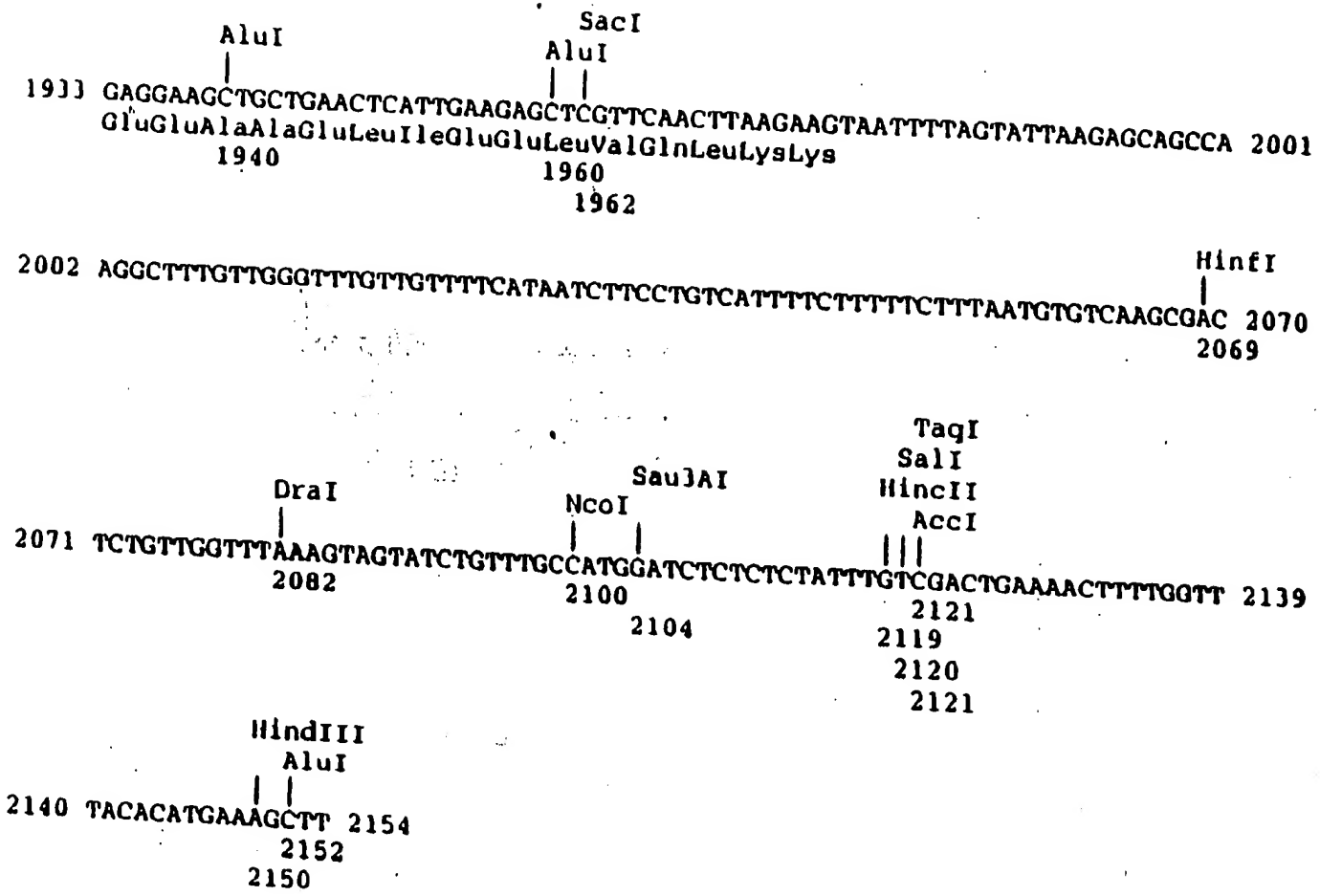
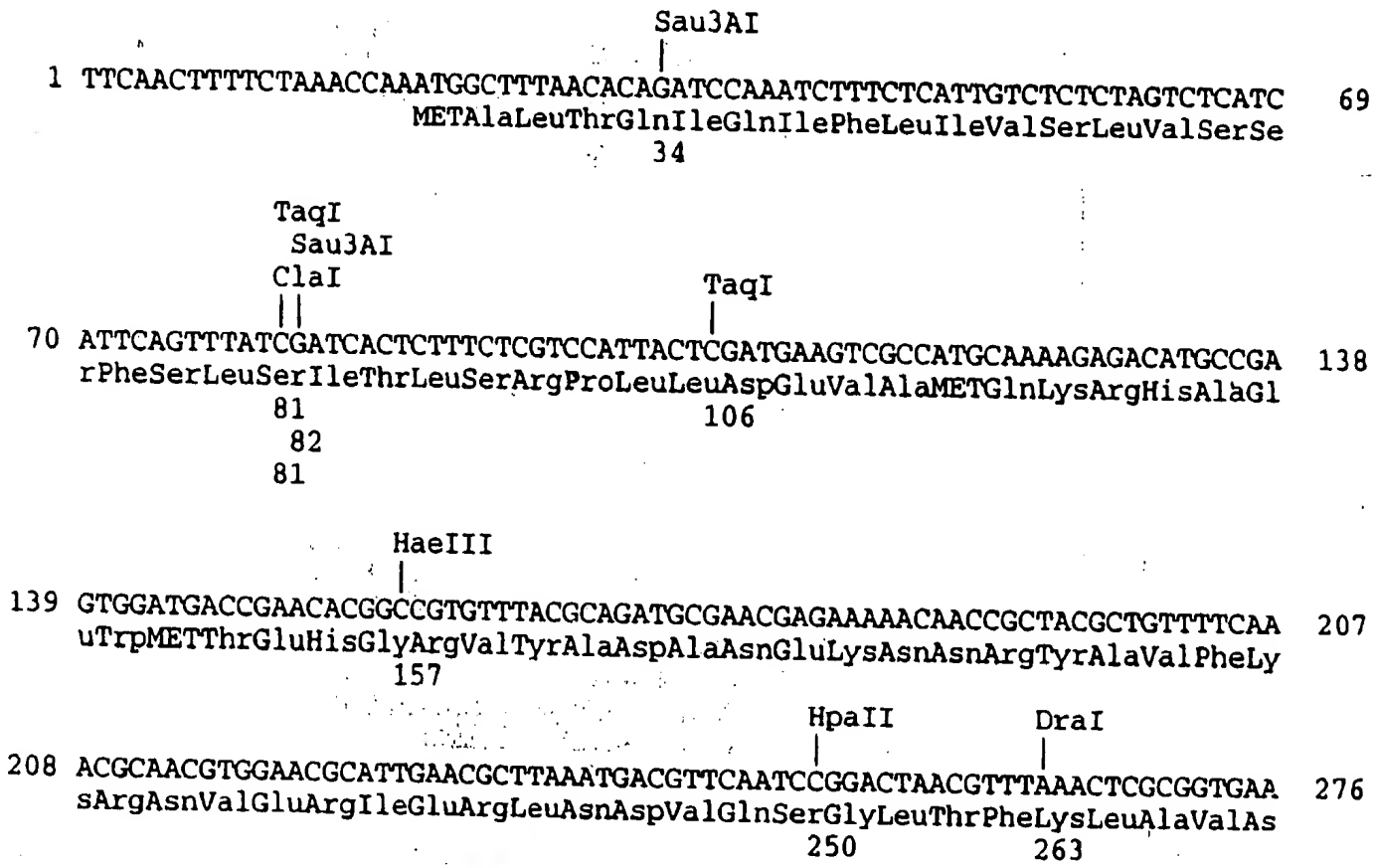

AluI
SacI
AluI
1933 GAGGAAGCTGCTGAACTCATTGAAGAGCTCGTTCAACTTAAGAAGTAATTTTAGTATTAAGAGCAGCCA 2001
GluGluAlaAlaGluLeuIleGluGluLeuValGlnLeuLysLys
1940
1960
1962
HinfI
2002 AGGCTTTGTTGGGTTTGTTGTTTTCATAATCTTCCTGTCATTTTCTTTTTCTTTAATGTGTCAAGCGAC 2070
2069
TaqI
SalI
DraI
SauJAI
HincII
NcoI
AccI
2071 TCTGTTGGTTTAAAGTAGTATCTGTTTGCCATGGATCTCTCTCTATTTGTCGACTGAAAACTTTTGGTT 2139
2082
2100
2104
2121
2119
2120
2121
HindIII
AluI
2140 TACACATGAAAGCTT 2154
2152
2150

FIG. 3G

FOREIGN DERIVED

Brassica Campestris Seed Specific cDNA-EA9



Complete nucleotide sequence of B. campestris cDNA EA9. The longest open reading frame is designated by three letter amino acid code. PolyA tails are evident at the end of the sequence and a potential polyadenylation signal is underlined.

FIG. 4A

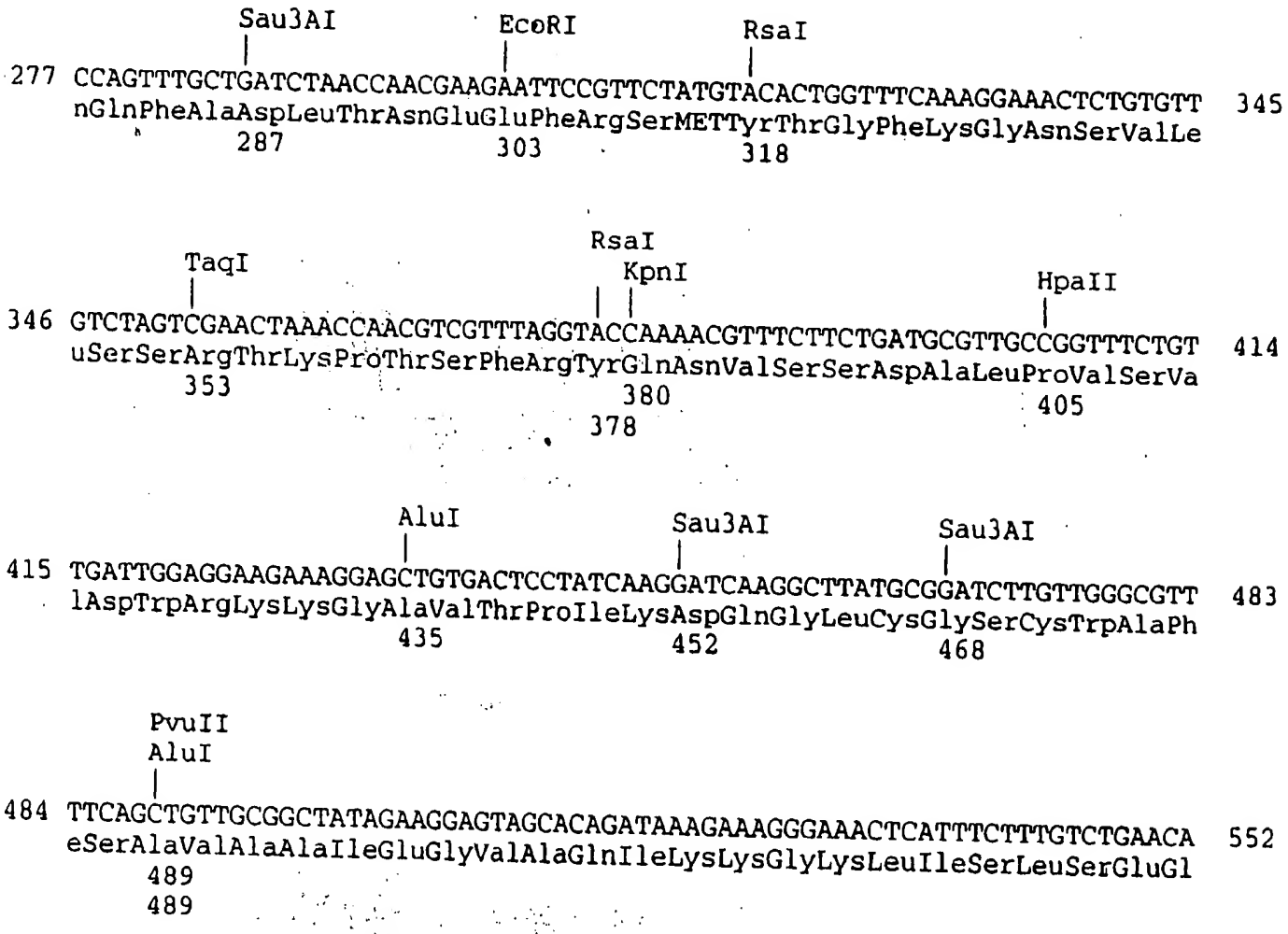

Sau3AI
EcoRI
RsaI
277 CCAGTTTGCTGATCTAACCAACGAAGAATTCCGTTCTATGTACACTGGTTTCAAAGGAAACTCTGTGTT 345
nGlnPheAlaAspLeuThrAsnGluGluPheArgSerMETTyrThrGlyPheLysGlyAsnSerValLe
287
303
318
TaqI
RsaI
KpnI
HpaII
346 GTCTAGTCGAACTAAACCAACGTCGTTTAGGTACCAAAACGTTTCTTCTGATGCGTTGCCGGTTTCTGT 414
uSerSerArgThrLysProThrSerPheArgTyrGlnAsnValSerSerAspAlaLeuProValSerVa
353
380
405
378
AluI
Sau3AI
Sau3AI
415 TGATTGGAGGAAGAAAGGAGCTGTGACTCCTATCAAGGATCAAGGCTTATGCGGATCTTGTTGGGCGTT 483
lAspTrpArgLysLysGlyAlaValThrProIleLysAspGlnGlyLeuCysGlySerCysTrpAlaPh
435
452
468
PvuII
AluI
484 TTCAGCTGTTGCGGCTATAGAAGGAGTAGCACAGATAAAGAAAGGGAAACTCATTTCTTTGTCTGAACA 552
eSerAlaValAlaAlaIleGluGlyValAlaGlnIleLysLysGlyLysLeuIleSerLeuSerGluGl
489
489

FIG. 4B

PCT/EP97/03250

FIG. 4C

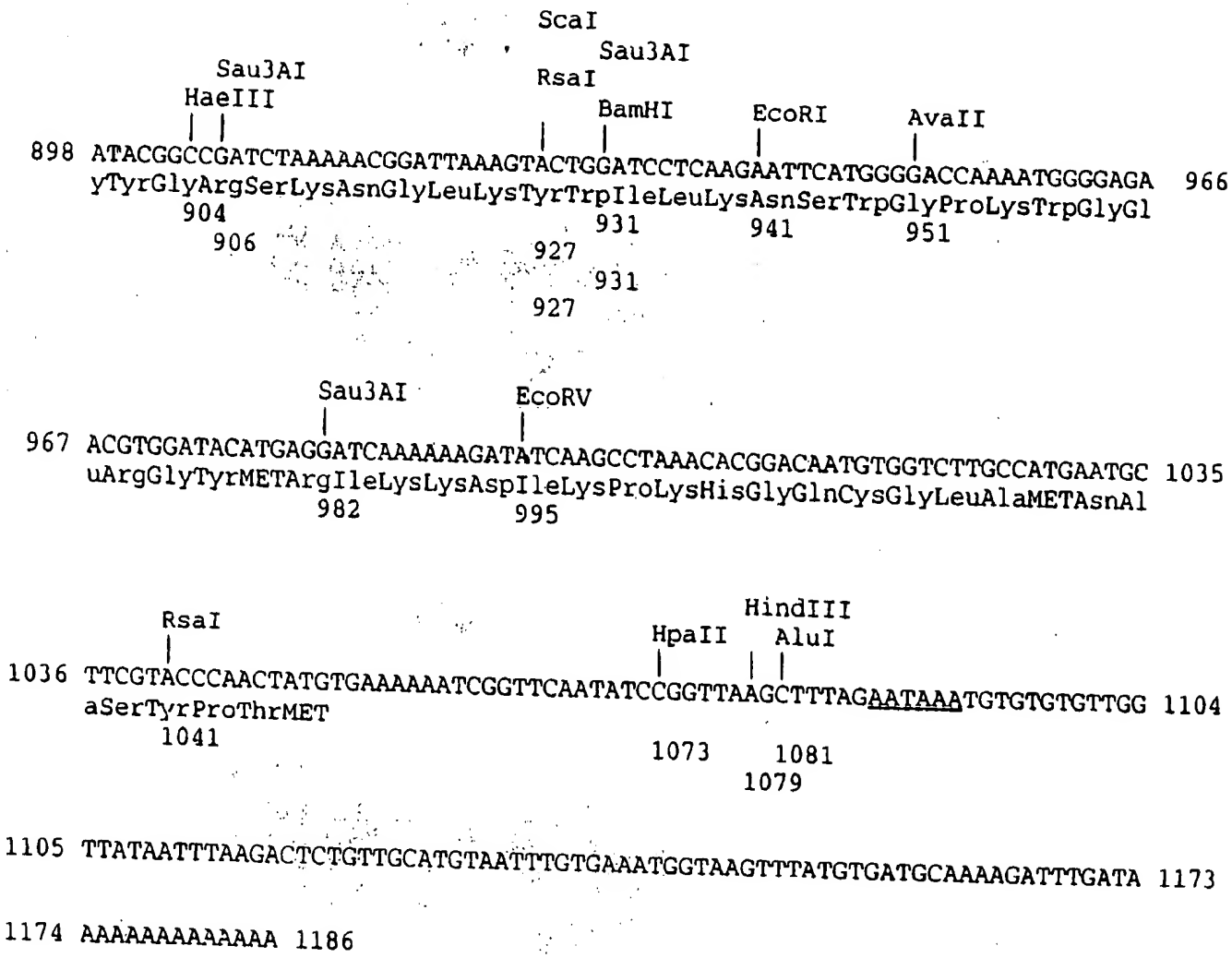

ScaI
Sau3AI
Sau3AI
RsaI
HaeIII
BamHI
EcoRI
AvaII
898 ATACGGCCGATCTAAAAACGGATTAAAGTACTGGATCCTCAAGAATTCATGGGGACCAAAATGGGGAGA 966
yTyrGlyArgSerLysAsnGlyLeuLysTyrTrpIleLeuLysAsnSerTrpGlyProLysTrpGlyGl
904
906
931
941
951
927
931
927
Sau3AI
EcoRV
967 ACGTGGATACATGAGGATCAAAAAAGATATCAAGCCTAAACACGGACAATGTGGTCTTGCCATGAATGC 1035
uArgGlyTyrMETArgIleLysLysAspIleLysProLysHisGlyGlnCysGlyLeuAlaMETAsnAl
982
995
RsaI
HindIII
HpaII
AluI
1036 TTCGTACCCAACTATGTGAAAAAATCGGTTCAATATCCGGTTAAGCTTTAGAATAAATGTGTGTGTTGG 1104
aSerTyrProThrMET
1041
1073
1081
1079
1105 TTATAATTTAAGACTCTGTTGCATGTAATTTGTGAAATGGTAAGTTTATGTGATGCAAAAGATTTGATA 1173
1174 AAAAAAAAAAAAA 1186

FIG. 4D

```
3H11 TTTTTTTGAGCAAAGGGCAACTCAGATATCCAAAGATGAATCCAACATATA 51

3H11 GCTTACAGCTGGGAGAACATTGTCTAACTCTTCTGAAATTTAAATGTTATC 102

3H11 CAGAATCCTTCATCATAAAATAATATCAAAATGCAAATCTATTTTTTCTAC 153

3H11 TCTTGTCTAGCTTCAACTTTCTTCTTCTGCTCATCAATTAGCAATTAATCC 204
                                    TGCTCATCAATTAGCAATTAATCC

3H11 AAAACCATTATGGCTGCCAAAAATTCAGAGATGAAGTTTGCTATCTTCTTC 255
2A11 AAAACCATTATGGCTGCCAAAAATTCAGAGATGAAGTTTGCTATCTTCTTC
           METAlaAlaLysAsnSerGluMETLysPheAlaIlePhePhe

3H11 GTTGTTCTTTTGACGACCACTTTAGTTGATATGTCTGGAATTTCGAAAATG 306
2A11 GTTGTTCTTTTGACGACCACTTTAGTTGATATGTCTGGAATTTCGAAAATG
     ValValLeuLeuThrThrThrLeuValAspMETSerGlyIleSerLysMET

3H11 CAAGTGATGGCTCTTCGAGACATACCCCCACAAGAAACATTGCTGAAAATG 357
2A11 CAAGTGATGGCTCTTCGAGACATACCCCCACAAGAAACATTGCTGAAAATG
     GlnValMETAlaLeuArgAspIleProProGlnGluThrLeuLeuLysMET

3H11 AAGCTACTTCCCACAAATATTTTGGGACTTTGTAACGAACCTTGCAGCTCA 408
2A11 AAGCTACTTCCCACAAATATTTTGGGACTTTGTAACGAACCTTGCAGCTCA
     LysLeuLeuProThrAsnIleLeuGlyLeuCysAsnGluProCysSerSer

3H11 AACTCTGATTGCATCGGAATTACCCTTTGCCAATTTTGTAAGGAGAAGACG 459
2A11 AACTCTGATTGCATCGGAATTACCCTTTGCCAATTTTGTAAGGAGAAGACG
     AsnSerAspCysIleGlyIleThrLeuCysGlnPheCysLysGluLysThr

3H11 GACCAGTATGGTTTAACATACCGTACATGCAACCTGTTGCCTTGAACAATA 510
2A11 GACCAGTATGGTTTAACATACCGTACATGCAACCTGTTGCCTTGAACAATA
     AspGlnTyrGlyLeuThrTyrArgThrCysAsnLeuLeuPro .
```

FIGURE 5A

```
3H11 TCAATGATCTATCGATCGATCTATCTATCTATTTATCTGTCTCTGCGCGTA 561
2A11 TCAATGATCTATCGATCGATCTATCTATCTATTTATCTGTCTCTGCGCGTA

3H11 TAGTGTTGTCTGTACCTTTGGTGTGAAGAATATGAATAAAGGGATACATAT 612
2A11 TAGTGTTGTCTGTACCTTTGGTGTGAAGAATATGAATAAAGGGATACATAT

3H11 ATCTAGATATATTCTAGGTAATGTCCTATTGTATTTAAAATTTGTAGCAAT 663
2A11 ATCTAGATATATTCTAGGTAATGTCCTATTGTATTTAAAATTTGTAGCAAT

3H11 GATTGTTTGAATAAAAACATACCATGAGTGAAATAATTATTCCACATTAAT 714
2A11 GATTGTTTGAATAAAAACATACCATGAGTGAAATAATTATTCC

3H11 TCACGTATTTATTTCACTTATGATACGTATTTTTGTTCCTTTCGCGTAAAA 765

3H11 AAAAAAAAA    774
```

FIGURE 5B

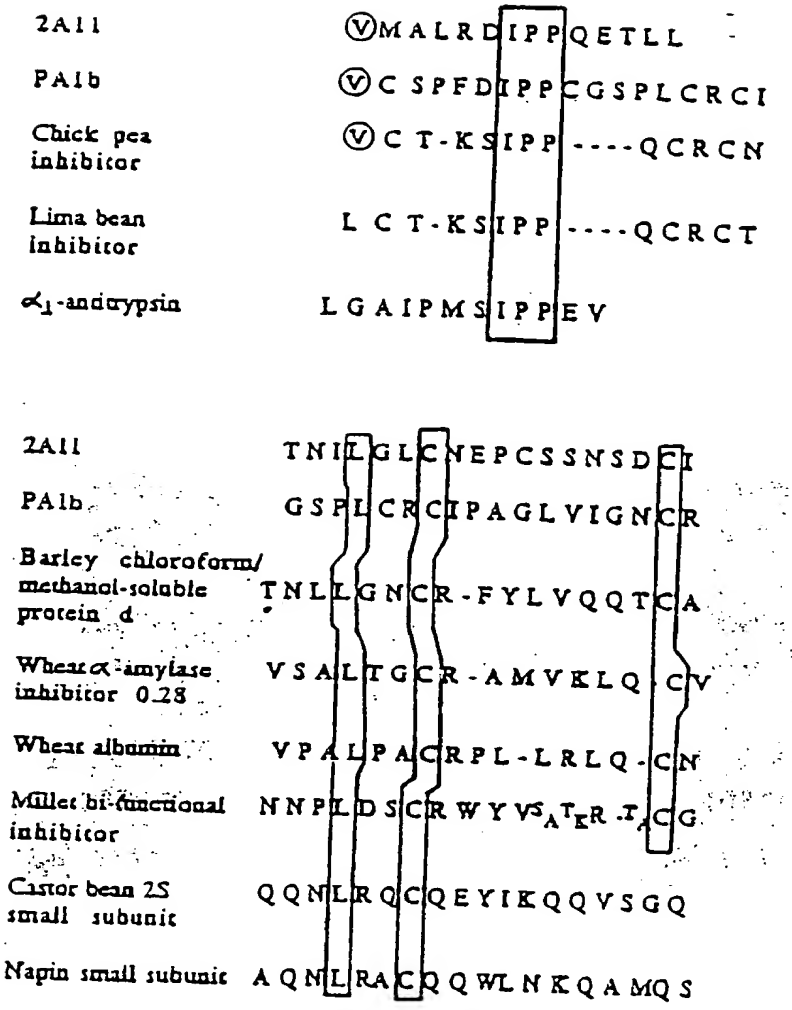

2A11
PA1b
Chick pea inhibitor
Lima bean inhibitor
α_1-antitrypsin
2A11
PA1b
Barley chloroform/ methanol-soluble protein d
Wheat α-amylase inhibitor 0.28
Wheat albumin
Millet bi-functional inhibitor
Castor bean 2S small subunit
Napin small subunit

FIGURE 6

2A11 GENOM.

```
        10          20          30          40          50
CTCGAGCCCT TTAAAAAGTA TAGTCAATAT TTACGGTGAC CGTGAATTTC
        60          70          80          90         100
TTAATTATGA TATATAATTT AAAAGAAATC ATGATCACAT TCTACTGATG
       110         120         130         140         150
AGAACATGTG CTAATCAAGG GAAAACATGG ATGTGAAAAA TACTTTTTGT
       160         170         180         190         200
TAAAAGTAAA AAAAAATGTG AAATTTTGTT AGTTATTTAC TACCTATACA
       210         220         230         240         250
TTATTTGAGC ATGTGCAAAC TTTACAAATA CCTAATAGAA GATTTTCACC
       260         270         280         290         300
TGCCTGTATA TATGTAAATT AATTATAATG AACACTCTCA CATAAAATAA
       310         320         330         340         350
TTATCAGTAT ATACATTAAT ACTTGCCCTC CACAATGAAT TAAATAAAAT
       360         370         380         390         400
GTAGAACATG ATCTACACTT CAATAAAACT AAGACCATAA AGAATAATTT
       410         420         430         440         450
CAAAATATAC ACATGTCAAC AATAAATTAT TTGCATATTA TATTAACTTA
       460         470         480         490         500
CTAAACAATC TTTACTTTTG AAATATAAAA ATAATCAAGT TATAAGTCTG
       510         520         530         540         550
CTCAAAGTAA AGCACTTGTT AGACTCATCT GATTTTGAGA AGGTAAGCAA
       560         570         580         590         600
ATTGATGGTG CATAATAGTC ACAAGTAAAA TATAAAATAG ATTTCATTAG
       610         620         630         640         650
TAAAATTGTT TTTTACTTTC TTTATATATA ATTATCAATA TCCTTCAATG
       660         670         680         690         700
GTAGGTTAAT TATATTGTTA ACTTCTTGTT GAATTAAAGC AATAAGACAA
       710         720         730         740         750
GAATATTAAA GATAAAAGAA CAATAAAAAT AGAAAGACTA AGAGATAAGA
       760         770         780         790         800
GTTTTCTTAT TCTTCTTTCA ATAAGTATCA TCAAGTGTAT ACAATATAAA
       810         820         830         840         850
TTTTTGTATT TTTGATCTAT CTATTTATAA TGTTATATAT AAGCATACAA
       860         870         880         890         900
AAGATCAGTC ATAAATATGA CTTTAATCAT GAAAATAATG AAAGAGATTA
       910         920         930         940         950
TGAAGGCGTA AGGTTACTAG AATAATAGTC ATTAAAAAAA GGGGTTATCT
       960         970         980         990        1000
TTATAATTGA ATAATTGATG AAGTAATGGA GATAATTAGT GAGCATAAAT
      1010        1020        1030        1040        1050
TTTTTTAAAA AAATGGACAT TTACACTATA ATATTTTATA ACACTTTCCC
      1060        1070        1080        1090        1100
TTAAACATCT AGGTATAAAT AATGAGTCTT GTCAAAATCT TAGTAGGAAA
```

FIGURE 7A

```
    1110        1120        1130         140         1150
AATTCTGTGA AATTTTTTTA GTGAAAACAA ATGATATAAA TATCTTGAAT
      1160        1170        1180        1190        1200
ACTCATTATT TGTTGTCTCA TTAAAAATCT TATCTGACCT ATAAAATAAA
      1210        1220        1230        1240        1250
TTATTTGCTC AACTCAAAAT AGTTTTTCAT TCTAAAATTA GTATAATTAT
      1260        1270        1280        1290        1300
TAGTGAATAT TTAATTAACA TAATTGTATA CTAAGGGGCC TATAAATTGG
      1310        1320        1330        1340        1350
ATTCTTCTCA AAGAAAAATA AAATCACCAC ACAACTTTCT TCTTCTGCTC
      1360        1370          1381        1390
ATCAATTAGC AATTAATCCA AAACCATT ATG GCT GCC AAA AAT
                                  MET Ala Ala Lys Asn
   1399        1408        1417        1426
TCA GAG ATG AAG TTT GCT ATC TTC TTC GTT GTT CTT TTG
Ser Glu MET Lys Phe Ala Ile Phe Phe Val Val Leu Leu
1435        1444        1454        1464        1474
ACG ACC ACT TTA GGTTCACAAC ACTTCTCCCT TATTTTGTTT
Thr Thr Thr Leu
      1484        1494        1504        1514        1524
TCTTAATTTC TTGGAAGTCA TATGCATGTG TTTGGTATCA TGGTATATAT
      1534        1544        1554        1564        1574
ATAAAGGAAA ATATTTTTCT TAATTACTGG TTTTCTAATG TTTGGTAGGT
      1584        1594        1604        1614        1624
AATCGGAAAT TATTATGAGA TAATGAACTT GCAAAGTCAT TATTATATAA
      1634        1644        1654        1664        1674
CTTTTTTTTT ATACTTTGAT TTAAGAATTC ATTTTTCTCA TTTTATATAA
      1684        1694        1704        1714        1724
ACTTATTTTT CAACAGAAAA TATTTTTCGA ACTATTCAAA CACACCCTAA
      1734        1744        1754        1764        1774
GACATTACAT ATATATATAT ATACACCCTC CGTTTTATAT TACTTAATGC
      1784        1794        1804        1814        1824
CTATTGAGTT GGCCCACCCT TTAAGAATGA TTCAATTAGA GATATGTTTT
      1834        1844        1854        1864        1874
ACTAAATTAA CCTATGCTTT AAGACTCTAA ATTTGGCTAT TACTATTTTA
      1884        1894        1904        1914        1924
CGTTGTAATT TAATGACAAA CATTTCATAA TGACTATAGT CTGAACTTAA
      1934        1944        1954        1964        1974
TTAGACAGAC GTATCTATAG TTTGCTTACT AATGATTCAT AGCTATATAT
      1984        1994        2004        2014        2024
TTGGAGAGGA GAGAGACAAA CGATATTAAG AAAGGGAGGA GAGAGGCGAG
      2034        2044        2054        2064        2074
GTAAATCTGA AATAGAGAAG AGAAAGGCAA CCAATTTTGA TCATCTATCA
      2084        2094        2104        2114        2124
TACTTTTGAT TATTATTTTT ATTATATGTA CGTTTACATT ACAGTTTTCG
```

FIGURE 7B

```
     2134        2144        2154         2164
AATTCTTACA TTAATCTTAA TCATAATATA TACA GTT GAT ATG
                                      Val Asp MET

   2173        2182        2191         2200
TCT GGA ATT TCG AAA ATG CAA GTG ATG GCT CTT CGA GAC
Ser Gly Ile Ser Lys MET Gln Val MET Ala Leu Arg Asp
2209        2218        2227         2236        2245
ATA CCC CCA CAA GAA ACA TTG CTG AAA ATG AAG CTA CTT
Ile Pro Pro Gln Glu Thr Leu Leu Lys MET Lys Leu Leu
       2254        2263        2272        2281
CCC ACA AAT ATT TTG GGA CTT TGT AAC GAA CCT TGC AGC
Pro Thr Asn Ile Leu Gly Leu Cys Asn Glu Pro Cys Ser
   2290        2299        2308         2317
TCA AAC TCT GAT TGC ATC GGA ATT ACC CTT TGC CAA TTT
Ser Asn Ser Asp Cys Ile Gly Ile Thr Leu Cys Gln Phe
2326        2335        2344         2353        2362
TGT AAG GAG AAG ACG GAC CAG TAT GGT TTA ACA TAC CGT
Cys Lys Glu Lys Thr Asp Gln Tyr Gly Leu Thr Tyr Arg
       2371        2380             2393        2403
ACA TGC AAC CTG TTG CCT TGA ACAATATCAA TGATCTATCG
Thr Cys Asn Leu Leu Pro  .
       2413        2423        2433        2443        2453
ATCGATCTAT CTATCTATTT ATCTGTCTCT GCGCGTATAG TGTTGTCTGT
       2463        2473        2483        2493        2503
ACCTTTGGTG TGAAGAATAT GAATAAAGGG ATACATATAT CTAGATATAT
       2513        2523        2533        2543        2553
TCTAGGTAAT GTCCTATTGT ATTTAAAATT TGTAGCAATG ATTGTTTGAA
       2563        2573        2583        2593        2603
TAAAAACATA CCATGAGTGA AATAATTATT CCACATTAAT TCACGTATTT
       2613        2623        2633        2643        2653
ATTTCACTTA TGATACGTAT TTTTGTTCCT TTCGCGTAGA TTTTTGATCC
       2663        2673        2683        2693        2703
TTTTCCCTTT TGAATATTAA ACATTAAACA CAAATAATGT TTATTAAATT
       2713        2723        2733        2743        2753
AAGTTAATAT TTTTATTTAG CTATTTATAT TTTTATTTGA AATCAAACTT
       2763        2773        2783        2793        2803
GATAAATATT TATAAAGATA ATTAACAAGT AATGTGACAC TAACACCATG
       2813        2823        2833        2843        2853
TAATATTATC TTGTCGTTAT TTATGATAAT ATTTTAAAAT TATAATTTCA
       2863        2873        2883        2893        2903
GTTAAAAAAT TATTAAAAAA ACATACTTTT AAAAAGTGAG TTAGCCTCCG
       2913        2923        2933        2943        2953
CTACCCACAT ACTTATGAAT TGGACTAGTT GTTTTTTGAC CCACAAAAAG
       2963        2973        2983        2993        3003
AATGGGCTAA TTAAACCTGA CCTATCAAAT TTCAGAATCT GCATAGATTA
```

FIGURE 7C

```
      3013        3023        3033        3043        3053
GTCCGAACGA AATGAGTCAG CCCGTATTGA ACAAAATATC AACAAGGACG
      3063        3073        3083        3093        3103
TTATGTAAAG ATGTTTAAGA AGGAAAAAAG ATTTCTAATA CATATGGACT
      3113        3123        3133        3143        3153
TTCAATATCC CAACTTTGTC TGGCGATCTG AACCCTGCTT AGTTTGTTGA
      3163        3173        3183        3193        3203
TCATTAACTT GTCTTGCTAT GTATTTAAGA TTTAAACTTT ATATGTTTAA
      3213        3223        3233        3243        3253
ACTTACAGAA AATACATATA AATCTCTCAA GACTTGGCAA CATAATTTAC
      3263        3273        3283        3293        3303
TTTAGTACTT AAACTACATG AAAATTTAAA TATCCTTTTA ACATCTTTGA
      3313        3323        3333        3343        3353
AGTGAATTAA ATTATCACAA TCCGAGCCTA CACCTTGGAC GTGGCCGGCA
      3363        3373        3383        3393        3403
CTCAAGAACC AGTGCTGGTC CCCAAGCTAA CCCTCATCCT GACTGACTAC
      3413        3423        3433        3443        3453
AAGCGGAAGG CTAACTTAAG TATACAAAAG CTTAAAACTG AATAAAATAA
      3463        3473        3483        3493        3503
ACTTTACAAG GTTTTAACAC AAATGAACAA CTTTGAAGAA AATAATATAT
      3513        3523        3533        3543        3553
TCAACTAGCC ATAAAATAGA CAACTTTAGT CTTTAAAACA TTTAATAAAA
      3563        3573        3583        3593        3603
TAAATGCAAA ATATAGACTC CTTAACTAAA CTGACTATCT ATGGAGCCTC
      3613        3623        3633        3643        3653
TAATTGATAA AGATGGAAGT CGGGACAAGA CCACGACATC CTGACTAAAC
      3663        3673        3683        3693        3703
TGAGAAGTAA ATAAAATCCC CCGGAAAAAA AGGAGCCTCA CCATGGCTAA
      3713        3723        3733        3743        3753
CTCGAACTCG GGGATATATC AATGAAGCTC CTGTTGATGA TCTTGAAGAC
      3763        3773        3783        3793        3803
ATGTCTCTGC ATCATCAAAA AGATGCAGGC CAAATGGCTC AGTACGTAAA
      3813        3823        3833        3843        3853
ATGTACGAGT ATGTAAGGGA AATTCTAAAG TATAACATAA GCTTGATACT
      3863        3873        3883        3893        3903
TGAATAAAAG GAAACATACT TACCTCTTTT CAACTCAACT CAAATTAAGA
      3913        3923        3933        3943        3953
ATAAGATACT CAACTCAAAG ATTAGGTATT CAACGCAAAT ATGGCACTCT
      3963        3973        3983        3993        4003
ACTCAATGAA GTACAAATTA ACTCAGGATA CTCGACTTAA GATACTCAAC
      4013        4023        4033        4043        4053
TCCCGACACT CAACTGAACT CATTTCAATA TAAAGCAGCT TAAAACAAGT
      4063        4073        4083        4093        4103
TCAGTATAAA GTAAAGTTGT TTAAAAACAT GATGTCAACT CTGTGTGTAT
      4113        4123        4133        4143        4153
AATAAGGGAT ACAACATAAC TTTGAAATGT ATATAAAAAT ACAATTAACT
```

FIGURE 7D

```
      416          4173         4183         4193         4203
GATGTATATA AAAATACATT AATCTATGGG AGATTCTCTA ACCGACAACC
      4213         4223         4233         4243         4253
ATCACTTAAG GGCTAAGATG ATGATATAGC GATCTACCGC ACGCTGCCAT
      4263         4273         4283         4293         4303
CGCATCTTAT ACCCGGCCAA AGGTATAAGA CCTGAACTGC CTAATGAATC
      4313         4323         4333         4343         4353
CACTAATAAA CTGTTAAAAG GAATCATCTA AAAAGTATGA CCCTTTTCTA
      4363         4373         4383         4393         4403
CCCATAGTGG CTAACATGGT TTATGGGGGC TGTGAGTTAT CTGAACTCTC
      4413         4423         4433         4443         4453
CCCCATATCG GTGCTCAATA CTACTCCAAA AAATATACTG CTCTTATGTT
      4463         4473         4483         4493         4503
TAAAAACATA CTGATTCTGT GGTTTGAAAT TATTGCTTAA AGCTTAGATT
      4513         4523         4533         4543         4553
TTTGAAAAGC TCTCTTTTGA AAATCGTAGT TTCCTTTTTC TTCTATTAAA
      4563         4573         4583         4593         4603
GCTAGACATA GGCTATGTAG AACTCTAGCT TACCTTCCTT CTCAAAAGTT
      4613         4623         4633         4643         4653
TGAAAACATT TGCTTAGATT CTTAGGGACT ACTTAGTTCC CTTGTTGGAA
TTC
```

FIGURE 7E

PG GENOMIC

```
        10          20          30          40          50
AAGCTTCTTA AAAAGGCAAA TTGATTAATT TGAAGTCAAA ATAATTAATT

        60          70          80          90          100
ATAACAGTGG TAAAGCACCT TAAGAAACCA TAGTTTGAAA GGTTACCAAT

        110         120         130         140         150
GCGCTATATA TTAATCAACT TGATAATATA AAAAAAATTT CAATTCGAAA

        160         170         180         190         200
AGGGCCTAAA ATATTCTCAA AGTATTCGAA ATGGTACAAA ACTACCATCC

        210         220         230         240         250
GTCCACCTAT TGACTCCAAA ATAAAATTAT TATCCACCTT TGAGTTTAAA

        260         270         280         290         300
ATTGACTACT TATATAACAA TTCTAAATTT AAACTATTTT AATACTTTTA

        310         320         330         340         350
AAAATACATG GCGTTCAAAT ATTTAATATA ATTTAATTTA TGAATATCAT

        360         370         380         390         400
TTATAAACCA ACCAACTACC AACTCATTAA TCATTAAATC CCACCCAAAT

        410         420         430         440         450
TCTACTATCA AAATTGTCCT AAACACTACT AAAACAAGAC GAAATTGTTC

        460         470         480         490         500
GAGTCCGAAT CGAAGCACCA ATCTAATTTA GGTTGAGCCG CATATTTAGG

        510         520         530         540         550
AGGACACTTT CAATAGTATT TTTTTCAAGC ATGAATTTGA AATTTAAGAT

        560         570         580         590         600
TAATGGTAAA GAAGTAGTAC ATCCCGAATT AATTCATGCC TTTTTTAAAT

        610         620         630         640         650
ATAATTATAT AAATATTTAT GATTTGTTTT AAATATTAAA ACTTGAATAT

        660         670         680         690         700
ATTATTTTTT TAAAAATTAT CTATTAAGTA CCATCACATA ATTGAGACGA

        710         720         730         740         750
AGGAATAATT AAGATGAACA TAGTGTTTAA TTAGTAATGG ATGGGTAGTA
```

FIGURE 8A

```
       760         770         780         790         800
AATTTATTTA TAAATTATAT CAATAAGTTA AATTATAACA AATATTTGAG

       810         820         830         840         850
CGCCATGTAT TTTAAAAAAT ATTAAATAGT TTGAATTTAA AACCGTTAGA

       860         870         880         890         900
TAAATGGTCA ATTTTGAACC CAAAAGTGGA TGAGAAGGGT ATTTTAGAGC

       910         920         930         940         950
CAATAGGRGG ATGAGAAGGA TATTTTGAAG CCAATATGTG ATGGATGAAG

       960         970         980         990        1000
GATAATTTTG TATCATTTCT AATACTTTAA AGATATTTTA GGTCATTTTC

      1010        1020        1030        1040        1050
CCTTCTTTAG TTTATAGACT ATAGTGTTAG TTCATCGAAT ATCATCTATT

      1060        1070        1080        1090        1100
ATTTCCGTCT TAAATTATTT TTTATTTTAT AAATTTTTTA AAAATAAATT

      1110        1120        1130        1140        1150
ATTTTTTCCA TTTAACTTTG ATTGTAATTA ATTTTTAAAA ATTACCAACA

      1160        1170        1180        1190        1200
TATAAATAAA ATTAATATTT AACAAAGAAT TGTAACATAA TATTTTTTTA

      1210        1220        1230        1240        1250
ATTATTCAAA ATAAATATTT TTAAACATCA TATAAAAGAA ATACGACAAA

      1260        1270        1280        1290        1300
AAAATTGAGA CGGGAGAAGA CAAGCCAGAC AAAAATGTCC AAGAAACTCT

      1310        1320        1330        1340        1350
TTCGTCTAAA TATCTCTCAT CCAAACTAAT ATAATACCCA TTATAATTAA

      1360        1370        1380        1390        1400
CCATATTGAC CAACTCAAAC CCCTTAAAAT CTATAAATAG ACAAACCCTT

      1410        1420        1430        1440        1450
CCCATACCTC TTATCATAAA AAAAATAATA ATCTTTTTCA ATAGACAAGT

      1460        1470        1480        1490        1500
TTAAAAACCA TACCATATAA CAATATATCA TGGTTATCCA AAGGAATAGT
```

FIGURE 8B

```
      1510        1520        1530        1540        1550
ATTCTCCTTC TCATTATTAT TTTTGCTTCA TCAATTTCAA CTTGTAGAAG

      1560        1570        1580        1590        1600
CAATGTTATT GATGACAATT TATTCAAACA AGTTTATGAT AATATTCTTG

      1610        1620        1630        1640        1650
AACAAGAATT TGCTCATGAT TTTCAAGCTT ATCTTTCTTA TTTGAGCAAA

      1660        1670        1680        1690        1700
AATATTGAAA GCAACAATAA TATTGACAAG GTTGATAAAA ATGGGATTAA

      1710        1720        1730        1740        1750
AGTGATTAAT GTACTTAGCT TTGGAGCTAA GGGTGATGGA AAAACATATG

      1760        1770        1780        1790        1800
ATAATATTGT AAGTATTTAA ATATTGGAAT ATATTTGTGG GGATGAAAAT

      1810        1820        1830        1840        1850
GATAGAGAAT ATAAGAATTA TTTGGAAGGA TGAAAAGTTA TATTTTATAA

      1860        1870        1880        1890        1900
AGTAGAAAAT TATTTTCTCG TTTTTAGTAA TTAAAGGTGA AAAATGAGTT

      1910        1920        1930        1940        1950
TTCTCGTAAG CGAGGAAAGT CATTTTCCAT GGAACTGTAT TTTTTTTTTA

      1960        1970        1980        1990        2000
CTTTTAATAA CGTCATAGTA TTTGCTATAC TCAAGAATAA GACACTATTA

      2010        2020        2030        2040        2050
TTGATGTTTA GTGCTCGAAA AGAAATTGAT AGTAATTTTG CTAATATAAC

      2060        2070        2080        2090        2100
TATCAATTTC TTATATGTAT ATTTTTCAAC CAAAATAACA AAGCGTAATC

      2110        2120        2130        2140        2150
CAATAAGTGG GCCTCTAGAA TAAAGAGTAA GTTCTATTAA TTCTTAACCT

      2160        2170        2180        2190        2200
TATTTAATTT TATGGAAACC TCGACAAAAC GACAATGCTC AACTTATATT

CGAATTC
```

FIGURE 8C

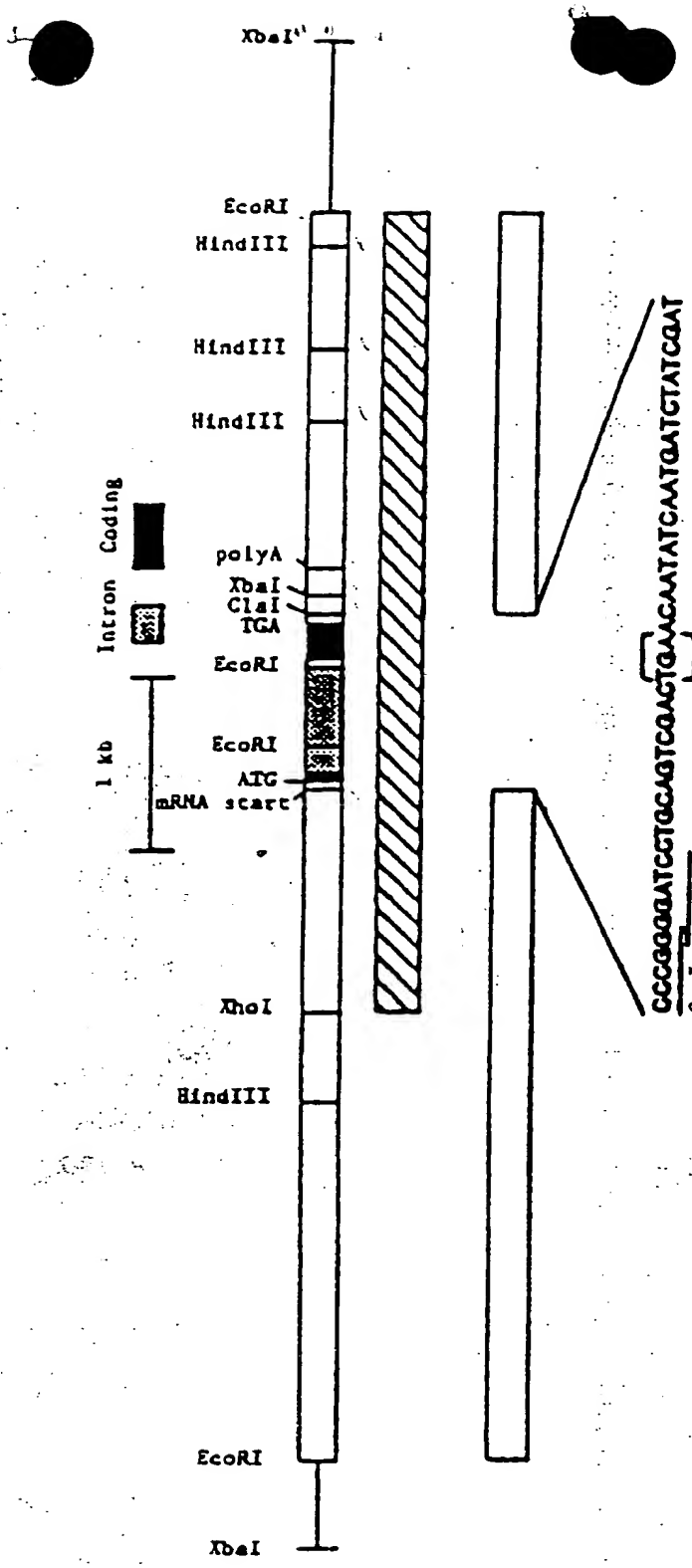

XbaI
EcoRI
HindIII
HindIII
HindIII
Intron Coding
polyA
XbaI
ClaI
TGA
EcoRI
1 kb
EcoRI
ATG
mRNA start
XhoI
HindIII
EcoRI
XbaI
SmaI
BamHI
PstI
SalI

FIGURE 9

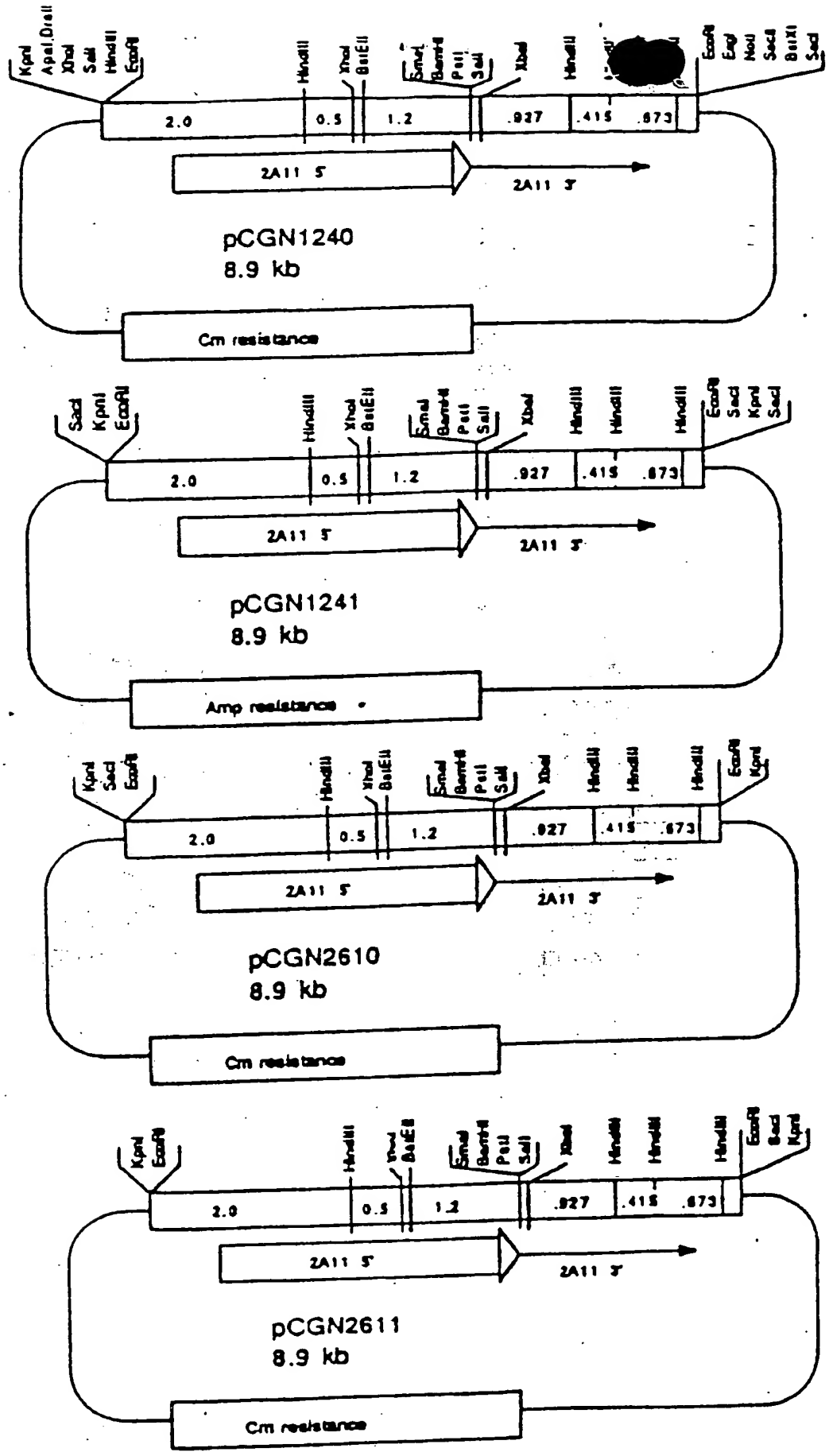

KpnI ApaI DraII XhoI SalI HindIII EcoRI
HindIII XhoI BstEII SmaI BamHI PstI SalI XbaI HindIII EcoRI EagI NotI SacII BstXI SacI
2.0
0.5
1.2
.927
.415
.873
2A11 5'
2A11 3'
pCGN1240
8.9 kb
Cm resistance
SacI KpnI EcoRI
HindIII XhoI BstEII SmaI BamHI PstI SalI XbaI HindIII HindIII HindIII EcoRI SacI KpnI SacI
2.0
0.5
1.2
.927
.415
.873
2A11 5'
2A11 3'
pCGN1241
8.9 kb
Amp resistance
KpnI SacI EcoRI
HindIII XhoI BstEII SmaI BamHI PstI SalI XbaI HindIII HindIII HindIII EcoRI KpnI
2.0
0.5
1.2
.927
.415
.873
2A11 5'
2A11 3'
pCGN2610
8.9 kb
Cm resistance
KpnI EcoRI
HindIII XhoI BstEII SmaI BamHI PstI SalI XbaI HindIII HindIII HindIII EcoRI SacI KpnI
2.0
0.5
1.2
.927
.415
.873
2A11 5'
2A11 3'
pCGN2611
8.9 kb
Cm resistance

FIGURE 10